

Algoritmen in het onderwijs

Een onderzoek in opdracht van het
College voor de Rechten van de Mens

Ed Smeets
Roos Geurts
Daniëlle van Helvoirt

met medewerking van Bente van Zijdam

Projectnummer: K2023076

© 2024 KBA Nijmegen / ResearchNed

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook, en evenmin in een retrieval systeem worden opgeslagen, zonder de voorafgaande schriftelijke toestemming van KBA Nijmegen / ResearchNed.

No part of this book/publication may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.

Samenvatting

Achtergrond van het onderzoek

Het College voor de Rechten van de Mens heeft KBA Nijmegen en ResearchNed verzocht om onderzoek te doen naar de risico's van de inzet van algoritmen in het onderwijs. De hoofdvraag voor het onderzoek luidt: *Welke risico's op discriminatie en uitsluiting (en kansen ter voorkoming hiervan) doen zich voor bij de inzet van algoritmen in het onderwijs in Nederland of kunnen zich in de toekomst voor gaan doen?*

Wat is en wat doet een algoritme?

Een algoritme is een procedure die gebruikt wordt om in een geautomatiseerd proces een probleem op te lossen of een berekening uit te voeren. Het is een lijst met stapsgewijs uit te voeren instructies. Algoritmen maken gebruik van data op basis waarvan geautomatiseerd beslissingen kunnen worden genomen of voorspellingen kunnen worden gedaan. Met behulp van vooraf ingevoerde data kunnen algoritmen worden 'getraind'. Gaandeweg worden zij dan tijdens het gebruik 'gevoed' met data die het algoritme verzamelt.

De aanpak van het onderzoek

Het onderzoek bestond uit deskresearch en het voeren van vraaggesprekken. In de fase van deskresearch zijn relevante organisaties aangeschreven en is onderzocht welk beleid zij voeren op het gebied van algoritmen in het onderwijs. Daarbij gaat het onder meer om sectororganisaties en andere organisaties die de belangen van onderwijsinstellingen behartigen en kennisorganisaties op het gebied van ICT in het onderwijs. Daarnaast is onderzoeksliteratuur gezocht en verwerkt (zowel uit Nederland als internationale publicaties). Er zijn drie vraaggesprekken gevoerd met experts en negen vraaggesprekken met 'stakeholders'. Daaronder waren vertegenwoordigers van leerlingen, studenten en ouders, ontwikkelaars en aanbieders van toepassingen, de Inspectie van het Onderwijs en het Ministerie van Onderwijs, Cultuur en Wetenschap. Het onderzoek had betrekking op het primair en voortgezet onderwijs, het middelbaar beroepsonderwijs en het hoger onderwijs.

Resultaten van het onderzoek

Toepassingen van algoritmen in het onderwijs

Algoritmen worden in alle onderwijssectoren gebruikt. In het primair en voortgezet onderwijs worden vooral adaptieve leersystemen gebruikt, evenals automatische nakijkprogramma's, dashboards, simulaties, leerlingvolgsystemen en adaptieve toetsen. In het mbo en hoger onderwijs wordt onder meer gebruik gemaakt van 'learning analytics', om de studievoortgang te volgen en/of de kwaliteit van het onderwijs te monitoren. In samenhang daarmee wordt in het mbo en het hoger onderwijs gewerkt aan modellen die in een vroeg stadium van de studie mogelijke studieuitval en de behoefte aan extra ondersteuning kunnen voorspellen. Ook zijn er in verschillende onderwijssectoren toepassingen gericht op het voorkomen van fraude en er zijn plaatsings-systemen die algoritmen gebruiken.

Risico's van het gebruik van algoritmen in het onderwijs

Zowel in de literatuur als in de gesprekken worden verschillende risico's van het gebruik van algoritmen in het onderwijs genoemd. Daarbij gaat het vooral om risico's op discriminatie en/of kansenongelijkheid. Dergelijke risico's kunnen het gevolg zijn van 'bias'. Deze kan ontstaan door de selectie of het gebruik van data waarmee het algoritme is getraind en door de manier waarop het algoritme is geprogrammeerd. Als het algoritme onvoldoende rekening houdt met de achtergrond of specifieke kenmerken van bepaalde groepen, kunnen zij worden benadeeld. Als voorbeeld worden culturele minderheidsgroepen genoemd, leerlingen die tweetalig zijn en leerlingen met beperkingen (zoals autisme, adhd of dyslexie). Na de inzet van het algoritme kunnen risico's ontstaan als de uitkomsten niet op bias worden gecontroleerd. Er wordt op gewezen dat hier een belangrijke taak ligt voor leraren die adaptieve leersystemen inzetten. Er wordt ook op gewezen dat het doorgaans moeilijk is om vast te stellen of bepaalde algoritmen inderdaad sommige groepen benadelen. Het is onvoldoende transparant hoe algoritmen precies werken. Volgens deskundigen komt dat mede doordat er commerciële belangen mee zijn gemoeid. Concrete voorbeelden van discriminatie of uitsluiting door de inzet van algoritmen in het onderwijs in Nederland heeft het onderzoek niet opgeleverd. Wel blijkt uit onderzoek in de Verenigde Staten dat detectoren die worden gebruikt om fraude bij het schrijven van teksten op te sporen, resultaten kunnen opleveren die nadelig zijn voor mensen voor wie Engels niet de eerste taal is.

Een ander risico van het gebruik van algoritmen in het onderwijs waarop wordt gewezen, is het risico dat de privacy van leerlingen en studenten onvoldoende is gewaarborgd. Zij worden doorgaans onvoldoende geïnformeerd over welke data worden verzameld en gebruikt. Ook speelt de kwestie van het eigenaarschap van de data. Er wordt op gewezen dat het niet wenselijk is dat de data bij een commerciële partij worden opgeslagen en dat de aanbieder van de toepassing eigenaar van de data is.

Tegenover deze risico's staan ook voordelen. Zo kunnen adaptieve leersystemen – mits zonder bias – eraan bijdragen dat leerlingen gedurende delen van de les in eigen tempo en op eigen niveau kunnen werken, waardoor de leraar werk uit handen wordt genomen. Ook kunnen goed functionerende algoritmen het signaal geven dat leerlingen of studenten extra ondersteuning of begeleiding nodig hebben om te voorkomen dat zij uitvallen. Tevens wordt erop gewezen dat docenten een subjectief beeld van studenten hebben dat zou kunnen worden genuanceerd door informatie die zij krijgen via een goed getraind en functionerend algoritme.

Beleid op het gebied van algoritmen in het onderwijs

Uit een rondgang bij betrokken organisaties en instanties blijkt dat er weinig beleid is op het gebied van algoritmen in het onderwijs. Sommige partijen zijn daar wel mee bezig. Met de Europese AI Act wordt gepoogd de risico's van artificiële intelligentie (en daarmee ook algoritmen) in te dammen, maar die regels vragen nog om een specifieke invulling. Er is een algoritmeregister van de Nederlandse overheid, waarmee wordt getracht tot meer transparantie te komen over toepassingen met algoritmen en hoe deze te werk gaan. Ook is er een referentiekader privacy en ethiek voor het hoger onderwijs en wordt een dergelijk kader ontwikkeld voor het mbo. Vertegenwoordigers van ouders en studenten/leerlingen merken dat betrokkenen slecht door scholen en andere onderwijsinstellingen worden geïnformeerd over toepassingen waarin algoritmen aan het werk zijn.

Wenselijk beleid op het gebied van algoritmen in het onderwijs

Deskundigen vinden het van groot belang dat wordt vastgelegd aan welke voorwaarden algoritmen moeten voldoen. Ook wordt het van groot belang geacht dat algoritmen uitgebreid worden getest op uitkomsten die tot kansongelijkheid (kunnen) leiden bij bepaalde groepen leerlingen en studenten. Ontwikkelaars en leveranciers dienen ervoor te zorgen dat aan de voorwaarden wordt voldaan en dat toepassingen voldoende getest worden. Daarnaast zou op landelijk niveau – door of namens de overheid, vertegenwoordigers van de onderwijssectoren of samenwerkingsverbanden die software inkopen – moeten worden onderzocht of toepassingen voldoen aan de eisen. Een aantal betrokkenen wijst erop dat onafhankelijke controle nodig is, gezien de belangen van de aanbieders. Als toepassingen vervolgens een keurmerk krijgen, is voor onderwijsinstellingen duidelijk dat deze betrouwbaar zijn. Onderwijsinstellingen hebben ook een verantwoordelijkheid bij de inkoop van toepassingen waarin algoritmen werkzaam zijn, maar zij beschikken over onvoldoende deskundigheid en mogelijkheden om deze uitgebreid te testen op onwenselijke uitkomsten en zij hebben onvoldoende invloed op de aanbieders van de toepassingen om aanpassingen te kunnen afdwingen.

Verder is van belang wie de eigenaar van de verzamelde data is: de leerling of student, de onderwijsinstelling of een externe (commerciële) partij. Ook is het van belang dat ouders, leerlingen en studenten worden geïnformeerd over de toepassing van algoritmen en dat hun toestemming wordt vastgelegd of dat er een mogelijkheid is dat leerlingen of studenten zich daaraan kunnen onttrekken ('opt out').

Gezien de snelle ontwikkelingen op het gebied van toepassingen van technologie in het onderwijs en de beperkte deskundigheid in het onderwijs op dat gebied, dienen onderwijsgevers en toekomstige onderwijsgevers te worden getraind in het werken met toepassingen waarin algoritmen worden gebruikt. Ook moeten zij de risico's daarvan (leren) onderkennen.

Inhoudsopgave

Samenvatting	iii
1 Inleiding	1
1.1 Aanleiding tot het onderzoek	1
1.2 Achtergrond	1
1.3 Begripsbepaling	2
1.4 Onderzoeksvragen	2
1.5 Onderzoeksopzet	3
1.6 Leeswijzer	4
2 Toepassingen van algoritmen in het onderwijs	5
2.1 Inleiding	5
2.2 Toepassingen van algoritmen in het onderwijs	5
3 Voordelen en risico's van algoritmen in het onderwijs	11
3.1 Inleiding	11
3.2 Kansen(on)gelijkheid	12
3.3 Vooroordelen of bias	13
3.4 Verlichting of belasting voor onderwijspersoneel	15
3.5 Privacy	16
4 Huidig en wenselijk beleid	17
4.1 Inleiding	17
4.2 Beleid van sectororganisaties en andere betrokken instanties	17
4.3 Welk beleid is wenselijk op het gebied van algoritmen in het onderwijs?	19
5 Conclusies	23
5.1 Inleiding	23
5.2 Conclusies	23
5.3 Beantwoording van de onderzoeksvragen	25
5.4 Aanbevelingen	26
Literatuurlijst	29

1 Inleiding

1.1 Aanleiding tot het onderzoek

Het College voor de Rechten van de Mens heeft KBA Nijmegen en ResearchNed verzocht om onderzoek te doen naar algoritmen in het onderwijs. Dit verzoek komt voort uit de wens van het College om risico's in kaart te brengen die het gebruik van algoritmen in het onderwijs met zich mee kan brengen op onder meer discriminatie en/of uitsluiting en in hoeverre algoritmen risico's opleveren voor kansengelijkheid. Het onderzoek naar algoritmen in het onderwijs dient tevens in beeld te brengen hoe eventuele risico's kunnen worden ondervangen, welke positieve effecten kunnen worden gerealiseerd door algoritmen in te zetten en wat het beleid ten aanzien van algoritmen is bij onderwijsinstellingen, sectororganisaties en andere betrokkenen.

1.2 Achtergrond

Uit een onderzoek dat het College liet uitvoeren naar de inzet van algoritmen bij sollicitatieprocedures en de mogelijke effecten daarvan, bleek dat algoritmen kansenongelijkheid in de hand kunnen werken bij het werven van personeel (Vervliet, 2022).

Kansengelijkheid in het onderwijs is een onderwerp waar veel onderzoek naar gedaan wordt. Het streven naar kansengelijkheid moet ertoe leiden dat ook leerlingen die onder minder gunstige omstandigheden opgroeien kansen krijgen om hun volledige potentieel te realiseren. Nederland heeft een lange traditie op het gebied van het bestrijden van onderwijsachterstanden. Sinds de jaren '70 voert de overheid gericht beleid om de verschillen in schoolloopbanen te reduceren die het gevolg zijn van een minder geprivilegieerde thuissituatie (Driessen, 2012). Uit onderzoek naar de ontwikkeling van de prestaties van de doelgroepen van het onderwijsachterstandenbeleid is gebleken dat tussen 1995 en 2011 een aanzienlijke vooruitgang is geboekt, met name bij de doelgroepleerlingen met een migratieachtergrond. De doelgroepleerlingen met een Nederlandse achtergrond liepen aanzienlijk minder van hun achterstand in. Uit hetzelfde onderzoek is gebleken dat de vooruitgang tussen 2003 en 2011 aanzienlijk minder groot was dan in de periode daarvoor en dat er de laatste jaren zelfs sprake was van een (lichte) achteruitgang (Driessen, 2012). In 2016 is het Actieplan Gelijke Kansen gepresenteerd (Ministerie van Onderwijs, Cultuur en Wetenschap, 2016). Hierin staan een aantal maatregelen om gelijke kansen in het onderwijs te bevorderen. Ook is in 2016 de basis gelegd voor de Gelijke Kansen Alliantie (GKA), wat sinds 1 oktober 2018 ook het Programma Gelijke Kansen is¹.

In een rapport van de OECD (2018) wordt geconcludeerd dat leerlingen met een migratieachtergrond vaak ondermaats presteren op school, met name als het migranten van de eerste generatie betreft. Sociaaleconomische achterstand en taalbarrières zijn twee van de grootste obstakels voor een succesvolle integratie van leerlingen met een migrantenachtergrond. Ook scoren zij gemiddeld lager op welbevinden dan leerlingen zonder migratieachtergrond.

1 Meer informatie hierover is te vinden op de website van de Gelijke Kansen Alliantie: <https://www.gelijke-kansen.nl/over-gelijke-kansen>

In 2019 is de Digitaliseringsagenda Primair en voortgezet onderwijs opgesteld (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019). In deze agenda komen verschillende aandachtspunten naar voren, waaronder de ethiek van digitalisering, digitale leermiddelen en innovatie. Ook de recent verschenen peiling Digitale Geletterdheid einde basisonderwijs (Inspectie van het Onderwijs, 2024) laat zien dat digitalisering hoog op de agenda staat.

De digitalisering in het onderwijs heeft invloed op de kansengelijkheid in het onderwijs (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019). Niet alle leerlingen hebben dezelfde kansen bij de digitalisering van het onderwijs, bijvoorbeeld doordat zij niet allemaal even digitaal geletterd zijn (Dondorp & Pijpers, 2020; Inspectie van het Onderwijs, 2024). Ook kan het zijn dat digitale toepassingen niet altijd aansluiten bij alle leerlingen (Dondorp & Pijpers, 2020). Uit de MYRA-monitor komt naar voren dat leraren, schoolleiders en bestuurders in het primair en voortgezet onderwijs gemiddeld aangeven dat zij het niet eens en niet oneens zijn met de stelling dat ICT helpt om kansenongelijkheid tegen te gaan (Karssen et al., 2023a, 2023b). Onderwijsgeevenden zien dus geen duidelijke voordelen van de inzet van ICT met als doel het realiseren van kansengelijkheid.

1.3 Begripsbepaling

Een algoritme is een procedure die gebruikt wordt om in een geautomatiseerd proces een probleem op te lossen of een berekening uit te voeren. Algoritmen fungeren als een exacte lijst van instructies die stap voor stap specifieke acties uitvoeren in hardware- of softwareroutines². Daarbij maken algoritmen gebruik van data op basis waarvan beslissingen kunnen worden genomen. Nadat hierop berekeningen zijn uitgevoerd, levert het algoritme bepaalde output, zoals het antwoord op een vraag of zoekterm. Algoritmen kunnen ook worden gebruikt om voorspellingen te doen, bijvoorbeeld over producten die consumenten waarschijnlijk willen kopen of over de kans dat studenten een opleiding zullen afronden. In het onderwijs kunnen algoritmen ook worden gebruikt om te voorzien in gepersonaliseerde leerroutes voor leerlingen die zelfstandig met educatieve software werken.

1.4 Onderzoeksvragen

Het College heeft voor dit onderzoek de volgende hoofdvraag gesteld: *Welke risico's op discriminatie en uitsluiting (en kansen ter voorkoming hiervan) doen zich voor bij de inzet van algoritmen in het onderwijs in Nederland of kunnen zich in de toekomst voor gaan doen?*

Op basis van deze hoofdvraag zijn de volgende deelvragen geformuleerd:

- 1) Hoe worden algoritmen binnen het onderwijs ingezet of wordt de inzet daarvan overwogen?
- 2) Welke risico's op discriminatie en/of uitsluiting door de inzet van algoritmen zijn gedetecteerd in het onderwijs in Nederland?
- 3) Welke risico's op discriminatie en/of uitsluiting door de inzet van algoritmen zijn niet gedetecteerd, maar kunnen verondersteld worden?
- 4) Welke risico's van de inzet van algoritmen zouden in de toekomst kunnen toenemen?

2 Deze omschrijving is te vinden op de website van het TechTarget Network, waar meer uitleg te vinden is over verschillende soorten algoritmen en hun werking: <https://www.techtarget.com/whatis/definition/algorithm>

- 5) In welke fase(n) en/of bij welke type(n) van onderwijs spelen deze risico's met name?
- 6) Wat is de stand van zaken van het beleid van het Ministerie van Onderwijs, de Inspectie van het Onderwijs en scholen op het gebied van het voorkomen van discriminatie en/of uitsluiting door de inzet van algoritmen in het onderwijs?
- 7) Wat kunnen het Ministerie van Onderwijs, de Inspectie van het Onderwijs en onderwijsinstellingen (meer, anders) doen om de (verkennde risico's op) discriminatie en/of uitsluiting door de inzet van algoritmen te voorkomen?

1.5 Onderzoeksopzet

Het onderzoek bestond uit twee hoofdfasen. Allereerst is gestart met een deskresearch, waarin relevante organisaties zijn aangeschreven en een websearch naar literatuur is uitgevoerd. Daarna is een reeks interviews uitgevoerd, eerst met experts en vervolgens met stakeholders. De onderzoeksfasen worden in deze paragraaf nader toegelicht.

Deskresearch

Aanschrijven van koepelorganisaties

In overleg met de opdrachtgever zijn alle onderwijskoepelorganisaties aangeschreven, te weten de PO-Raad, VO-raad, MBO Raad, Vereniging Hogescholen en de Universiteiten van Nederland. Gelijktijdig werden ook twee relevante stakeholders in het veld benaderd. Dit waren Kennisnet en het Nationaal Onderwijslab AI. Een aantal koepelorganisaties reageerden dat zij zelf geen beleid hadden, maar wezen ons wel op andere stakeholders die mogelijk informatie hadden. Daardoor zijn de Onderwijsraad, SURF, Npuls, de Inspectie van het Onderwijs, SIVON en NL AIC aangeschreven. Niet alle aangeschreven stakeholders hebben informatie kunnen of willen aanleveren aan de onderzoekers. Sommigen verwezen door naar anderen.

Literatuursearch

Naast het aanschrijven van de koepelorganisaties en andere relevante organisaties hebben we op verschillende manieren gezocht naar wetenschappelijke artikelen over het onderwerp. In de eerste plaats is hierbij gebruik gemaakt van de zoekmachine 'Web of Science'.

Het zoeken naar literatuur via 'Web of Science' met de zoektermen 'education AND (algorithms NEAR) risk' in Abstracts, gepubliceerd in de afgelopen vijf jaar, in de categorie 'Educational Research', met als documenttype 'Article' of 'Review Article', levert 35 artikelen op. Het merendeel heeft betrekking op het voorspellen van studiesucces of -uitval in het hoger onderwijs. Hierbij wordt vaak nagegaan welk type algoritme (vaak betiteld als 'machine learning models') de beste voorspelling levert en er wordt nauwelijks ingegaan op risico's van de inzet van algoritmen. De zoekterm 'risk' leidt overigens tot publicaties die betrekking hebben op 'students at risk' en niet tot publicaties over risico's van het gebruik van algoritmen. Vervolgens is ook gezocht met de zoektermen '(education OR school) AND (algorithms OR machine learning) AND (discrimination OR exclusion)'. Dit leverde echter geen aanvullende relevante publicaties op. Via 'Google' en 'Google Scholar' zijn ook enkele (Engelstalige en Nederlandstalige) relevante publicaties gevonden. Door het gebrek aan mogelijkheden tot filtering zijn deze zoekmachines echter minder geschikt voor een systematische literatuursearch.

In aanvulling op de literatuursearch via zoekmachines zijn recente jaargangen doorgenomen van de wetenschappelijke tijdschriften 'Computers and Education Artificial Intelligence' en 'International Journal of Artificial Intelligence in Education'. Het betreft de jaargangen van 2021 tot 2024. Daarnaast is literatuur verwerkt waarop we zijn geattendeerd door de opdrachtgever en door een aantal deelnemers aan de vraaggesprekken. Ook zijn aanvullende publicaties gevonden via de referenties in wetenschappelijke artikelen ('snowballing').

Interviews

Om, na de deskresearch, meer zicht te krijgen op de mogelijkheden en risico's van het gebruik van algoritmen in het onderwijs, zijn interviews uitgevoerd. Hierbij is gekozen om zowel met experts als met stakeholders te spreken. De experts zijn gekozen vanwege expertise op het gebied van algoritmen en hun toepassing in het onderwijs. De geïnterviewde experts zijn van de volgende organisaties: Kennisnet, Tilburg University en ROC Amsterdam – Flevoland / Hogeschool Windesheim.

Nadat er met de experts is gesproken, is ook met stakeholders in het onderwijs gesproken. Deze stakeholders zijn betrokken bij verschillende schakels van het onderwijs. Zo is er bijvoorbeeld gesproken met leerling- en studentvertegenwoordigers, een vertegenwoordiger van ouders, methodeontwikkelaars en toezichthouders. De stakeholders waarmee gesprekken zijn gevoerd, zijn:

- Bureau ICE
- CitoLab
- Inspectie van het Onderwijs
- ISO
- JOB
- LAKS
- Malmberg
- Ministerie van Onderwijs, Cultuur en Wetenschap (hierna OCW)
- Ouders & Onderwijs

Sommige stakeholders die zijn benaderd met het verzoek aan een vraaggesprek mee te werken, hebben niet op dat verzoek gereageerd, hebben aangegeven dat zij niet konden meewerken of hebben aangegeven dat zij niet met algoritmen werken.

In de oorspronkelijke opzet van het onderzoek zou er na de deskresearch en de gesprekken met experts gekozen worden om meer focus aan te brengen in de onderzoeksopzet door te kiezen voor één of enkele onderwijssectoren. Door de brede informatie die gevonden is, is er echter voor gekozen om het onderzoek niet in die zin te versmallen en ook breed stakeholders te bevragen.

1.6 Leeswijzer

In de volgende drie hoofdstukken worden de onderzoeksresultaten beschreven. Hierbij worden drie hoofdthema's onderscheiden:

- 1) Toepassingen van algoritmen in het onderwijs;
- 2) Voordelen en risico's van algoritmen in het onderwijs;
- 3) Huidig en wenselijk beleid op het gebied van algoritmen in het onderwijs.

Deze thema's komen achtereenvolgens aan bod in hoofdstuk 2, 3 en 4. In hoofdstuk 5 worden de belangrijkste conclusies besproken en worden de onderzoeksvragen beantwoord.

2 Toepassingen van algoritmen in het onderwijs

2.1 Inleiding

In dit hoofdstuk beschrijven we verschillende toepassingen van algoritmen in het onderwijs. Dat gebeurt op basis van de verkenning in het kader van de deskresearch, de literatuurstudie en de vraaggesprekken met experts en met stakeholders. Waar mogelijk en zinvol worden de resultaten uitgesplitst naar onderwijssector.

2.2 Toepassingen van algoritmen in het onderwijs

Het gebruik van algoritmen verschilt per onderwijssector. Deze verschillen kunnen gaan over het soort algoritmen dat worden gebruikt, de hoeveelheid algoritmen die wordt gebruikt en de redenering achter het gebruik van algoritmen. Daarom is er in deze rapportage gekozen om, indien mogelijk, per onderwijssector te rapporteren.

Een van de vormen van algoritmen die in het onderwijs worden gebruikt, is 'learning analytics' (Autoriteit Persoonsgegevens, 2023a). Dit is het gebruik van data om inzicht te krijgen in de voortgang van leerlingen en studenten en in de kwaliteit van het onderwijs. Er zijn twee toepassingsdoeleinden van 'learning analytics'. Toepassingen kunnen zijn gericht op het verbeteren van begeleiding en doorstroom van studenten en toepassingen kunnen zijn gericht op het verbeteren van het onderwijs. Vooral in het mbo en in het hoger onderwijs wordt gebruik gemaakt van 'learning analytics'.

Sectoroverstijgend

Medewerkers van SURF werken aan een verkenning naar een algoritme-register in het onderwijs in samenwerking met Npuls. Zij gaven aan dat zij weinig documenten en onderzoeksresultaten over dit onderwerp kunnen vinden. Ten behoeve van het register houden zij onder andere bijeenkomsten met mbo-, hbo- en wo-instellingen.

Daarnaast gebruikt de Inspectie van het Onderwijs algoritmen. Deze worden ingezet vanuit de toezichtrol die de Inspectie heeft.³ Zo maakt de Inspectie gebruik van een algoritme om een risico-inschatting te maken voor scholen in alle onderwijssectoren. Het algoritme is niet zelflerend, maar gebruikt vooraf vastgestelde, geautomatiseerde beslisregels. Vervolgens kijken inspecteurs en analisten of het nodig is om onderzoek te doen naar de uitslag van dit algoritme. Voordat er een toezichtactiviteit (bijvoorbeeld een inspectieonderzoek) wordt uitgezet, wordt er een menselijke beoordeling gegeven. Dit algoritme staat ook geregistreerd in het Algoritmeregister van de Nederlandse Overheid.⁴

3 De werkwijze van de Inspectie van het Onderwijs bij deze Prestatiemonitor is beschreven op de website van de inspectie: <https://www.onderwijsinspectie.nl/onderwerpen/werkwijze-van-de-inspectie/jaarlijkse-prestatieanalyse>.

4 Het doel van dit algoritme, wat het doet en welke data worden gebruikt, is beschreven op de website van het Algoritmeregister: <https://algoritmes.overheid.nl/nl/algoritme/79978718>

De experts en stakeholders geven aan dat zelflerende algoritmen op dit moment (nog) niet in de praktijk worden gebruikt. Er vinden wel pilots plaats, maar deze ontwikkeling staat nog in de kinderschoenen en heeft nog een lange weg te gaan. Daarnaast is het ook mogelijk dat zelflerende algoritmen wel al worden toegepast, maar dat het onbekend is waar, aangezien AI vaak niet open source is.

Funderend onderwijs

Binnen het funderend onderwijs worden verschillende algoritmen gebruikt zoals adaptieve leer-systemen, automatische nakijkprogramma's, dashboards en simulaties (Onderwijsraad, 2022). Uit een onderzoek in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap is gebleken dat zo'n veertig procent van de leraren in het funderend onderwijs behoefte heeft aan onafhankelijk advies over het gebruik van digitale leermiddelen (van Dam, Notermans & Vlieg, 2022).

Adaptieve leersystemen selecteren oefenstof voor individuele leerlingen op basis van hoe goed een leerling eerdere oefeningen heeft gemaakt. Deze leersystemen spelen een grote rol in het primair onderwijs en in mindere mate in het voortgezet onderwijs (Autoriteit Persoonsgegevens, 2023b). Volgens het Rathenau Instituut (2022b) werkt naar schatting 60 procent van de basisscholen met adaptief leermateriaal (maar er wordt niet aangegeven waarop die schatting is gebaseerd). In hetzelfde rapport wordt aangegeven dat Snappet en Gynzy de meest ingezette systemen zijn op basisscholen; daarnaast worden Rekentuin, Taalzee, Muiswerk, Words&Birds en Squla genoemd als adaptieve leersystemen die op basisscholen worden gebruikt. In 2020 hadden de, destijds, vier grootste aanbieders van adaptieve leersystemen (Squla, Gynzy, Snappet en Rekentuin) samen ongeveer 1.8 miljoen gebruikers (Bureau Common Ground & it's public, 2020). In het voortgezet onderwijs worden adaptieve leersystemen volgens LAKS met name in de onderbouw ingezet (Autoriteit Persoonsgegevens, 2023b).

Uit de MYRA-monitor, de Monitor digitalisering funderend onderwijs, blijkt dat in het primair onderwijs 93 procent van de leraren digitaal leermateriaal gebruikt dat bij de lesmethode hoort; 63 procent gebruikt dashboards waarop de voortgang van leerlingen te volgen is (Snappet, Gynzy, Prowise Learn, Rekentuin) en 40 procent gebruikt adaptieve gepersonaliseerde digitale leermiddelen. In het voortgezet onderwijs zijn die percentages respectievelijk 87, 31 en 29. Gemiddeld genomen zijn respondenten (leraren, bestuurders, schoolleiders) van mening dat de visie van de school of het schoolbestuur op digitale leermiddelen niet duidelijk is. Hun eigen digitale vaardigheden schatten zij in het primair onderwijs in het algemeen redelijk hoog en in het voortgezet onderwijs hoog in. Hun kennis over artificiële intelligentie schatten zij laag in (Karssen et al., 2023a, 2023b).

Algoritmen kunnen ook worden toegepast op onderwijsprocessen. Zo kunnen algoritmen worden toegepast bij wettelijke regels rondom plaatsing van leerlingen (Streefkerk & van Putten, 2020). Het gebruik van dit soort algoritmen kan leiden tot gevoelens van ongelijkheid bij ouders, schrijven de auteurs. In één van de interviews met stakeholders wordt opgemerkt dat algoritmen in systemen voor plaatsing van leerlingen op scholen ook discriminerend kunnen werken zonder dat ouders zich daarvan bewust zijn. Dat kan bijvoorbeeld het geval zijn als leerlingen uit een bepaalde wijk voorrang krijgen omdat de school in die wijk staat terwijl daar vooral ouders met hogere inkomens wonen. Kasman en Valant (2019) noemen enkele praktische risico's van algoritmen voor plaatsing van leerlingen die in de Verenigde Staten worden gebruikt: ouders die te weinig voorkeursscholen invullen en daardoor na de eerste ronde geen plek hebben, en ouders

die proberen het systeem te misleiden. Het gegeven dat niet alle leerlingen op de school van voorkeur kunnen worden geplaatst, kan tot negatieve percepties ten aanzien van het plaatsings-systeem leiden. Daar staat volgens de onderzoekers tegenover dat deze systemen eenvoudiger en efficiënter zijn dan een systeem van aanmelding bij scholen zelf en dat de kans groter is dat een leerling op de school van voorkeur wordt geplaatst.

Daarnaast gebruikt de Inspectie van het Onderwijs in het basisonderwijs een gestandaardiseerde maat om te kijken of leerlingen genoeg hebben geleerd in een schooljaar.⁵ Hierbij wordt gekeken naar de referentieniveaus en wordt rekening gehouden met de schoolweging. Ook dit is geen zelflerend algoritme, maar bestaat uit geautomatiseerde beslisregels.

5 Meer informatie hierover is te vinden op de website van de Inspectie van het Onderwijs: <https://www.onderwijsinspectie.nl/onderwerpen/onderwijsresultaten-primair-onderwijs/naar-een-nieuw-onderwijsresultatenmodel>.

Een voorbeeld van een AI-toepassing in het funderend onderwijs is Perusall (SURF, 2022). Perusall is een applicatie die door leerkrachten wordt gebruikt om leerlingen te ondersteunen en te motiveren bij het voorbereiden van klassikale lezingen. Leerkrachten kunnen online stof aanbieden via Perusall, waardoor leerlingen worden aangemoedigd om te gaan lezen. Bij de applicatie leiden diepgaandere aantekeningen van leerlingen tot een betere score. Het algoritme in Perusall is een geragsveranderaar die opmerkingen en *nudges* (een 'duwtje in de rug') voor studenten verzamelt om slimme feedback te geven. Perusall gaat uit van de intrinsieke motivatie van studenten en de externe motivatie van automatische beoordeling en gaat ervan uit dat studenten gemotiveerd zijn en goede cijfers willen behalen.

Het is niet altijd duidelijk of er in toepassingen in het onderwijs algoritmen aan het werk zijn en of dat risico's met zich meebrengt. Bij Ouders & Onderwijs vindt men dat kritischer moet worden gekeken naar toepassingen die vanuit de (lokale) overheid naar het onderwijs komen. Daarbij gaat het bijvoorbeeld om zogenoemde kansencirkels of monitoren waarvan data gekoppeld kunnen worden, zodat daaruit conclusies kunnen worden getrokken over kinderen en ouders. Ouders stellen bijvoorbeeld vragen over de GGD-monitor die leerlingen verplicht invullen op school. Zit daar een voorspellende waarde in en hoe werkt dat, zo vragen sommige ouders zich af. Op basis van de antwoorden die door leerlingen in het tweede en vierde leerjaar voortgezet onderwijs worden ingevuld in de basisvragenlijst Gezondheidsmonitor Jeugd⁶ wordt bijvoorbeeld bepaald welk percentage jongeren een 'afwijkende score' of een score in het 'grensgebied' heeft op psychosociale gezondheid. Ook wordt bepaald welk percentage voldoende weerbaar is. Verder wordt onder meer gevraagd of de ouders van de leerling moeite hebben om financieel rond te komen.⁷

Mbo

Zoals hierboven beschreven, spelen 'learning analytics' een grote rol in het mbo. Eegdeman et al. (2022) deden onderzoek naar het met behulp van artificiële intelligentie (AI) voorspellen van studiesucces in het middelbaar beroepsonderwijs in Nederland. Bij dit onderzoek werden 9 docenten betrokken en 95 eerstejaarsstudenten van de sportacademie. Aan het begin van het programma voorspelden docenten studieuitval beter dan de drie algoritmen die werden gebruikt; dit met name doordat docenten beter in staat waren om studenten te identificeren met een zeer grote kans op uitval dan de algoritmen konden. Na de eerste periode nam de nauwkeurigheid van de voorspelling zowel bij de docenten als bij de algoritmen toe, maar de algoritmen gingen beter voorspellen dan de docenten. Eegdeman (2023) geeft aan dat drie databronnen werden gebruikt om het model (met de drie algoritmen) te voeden:

- 1) data uit een intake-test gericht op cognitie en persoonlijkheid die studenten invulden;
- 2) data uit een vragenlijst die studenten invulden over hun verwachtingen;
- 3) data over achtergrondkenmerken van studenten, studieprestaties en gegevens over verzuim.

In totaal werden 142 variabelen opgenomen om te voorspellen welke studenten meer kans liepen om uit te vallen en dus in aanmerking kwamen voor extra begeleiding. In dit onderzoek is niet nagegaan in hoeverre er een risico was op discriminatie en uitsluiting door de inzet van algoritmen. De onderzoekers stellen juist dat leraren factoren die ten grondslag liggen aan studieuitval observeren die niet beschikbaar zijn voor algoritmen. Daarmee zouden zij uitval aanvankelijk

6 De vragenlijst van de gezondheidsmonitor kan worden geraadpleegd op een website van gezondheidsmonitoren: <https://www.monitorgezondheid.nl/documenten/basisvragenlijst-gezondheidsmonitor-jeugd-2019>

7 Meer informatie hierover is te vinden op de website met statistische gegevens van het Rijksinstituut voor Volksgezondheid en Milieu: <https://statline.rivm.nl/#/RIVM/nl/dataset/50077NED/table?ts=1700054504952>

beter kunnen voorspellen dan algoritmen, maar de onderzoekers wijzen ook op het gevaar dat de subjectieve observaties van leraren bevooroordeeld en selectief kunnen zijn, wat ten koste kan gaan van de voorspellende kracht. Op basis van de uitkomsten van het onderzoek stellen de onderzoekers dat algoritmen leraren kunnen ondersteunen bij het nemen van beslissingen over extra ondersteuning aan studenten. In aanvulling hierop toonden Eegdeman et al. (2023) aan dat het periodiek toevoegen van informatie aan modellen die studieuitval voorspellen tot betere voorspellingen leidt. Door die periodieke toevoeging ging het model beter voorspellen dan de docenten.

Tussen mbo-instellingen bestaan verschillen in welke data wel of niet worden gebruikt door dergelijke systemen. JOB geeft hierbij aan signalen te krijgen van scholen die bijvoorbeeld algoritmen gebruiken om te voorspellen of een student een vergrote kans heeft op het niet afmaken van een opleiding. Indicatoren die dan gebruikt worden, zijn de behaalde cijfers, aanwezigheid, maar ook bijvoorbeeld leeftijd en woonlocatie. Ook geeft JOB aan dat ze signalen ontvangen dat mbo-instellingen elk weer hun eigen indicatoren gebruiken voor deze algoritmen. Studenten zijn zich vaak niet bewust van het gebruik van dergelijke algoritmen. Ze weten vaak nog wel dat hun resultaten en aanwezigheid worden bijgehouden, maar dat persoonlijke data (bijv. woonlocatie, vooropleiding) ook wordt gebruikt, is onbekend voor studenten. Studenten krijgen niet te maken met automatische besluitvorming als gevolg van een algoritme dat bijvoorbeeld studieuitval voorspelt. Wel worden ze aan de hand van de uitslag van een dergelijk algoritme uitgenodigd voor een gesprek met hun mentor.

Een van de mogelijke voordelen van een dergelijk systeem is dat het leerlingen/studenten eruit haalt die anders niet waren opgemerkt. Echter wordt dit in de praktijk sterk betwijfeld door JOB, aangezien leerlingen/studenten al worden beoordeeld door een team van docenten. Ook LAKS geeft aan dat juist het risico bestaat dat leerlingen buiten de boot vallen als te veel wordt gevaren op een leerlingvolgsysteem. Een dergelijk systeem neemt namelijk geen kenmerken als stress, levensgeluk en ontwikkeling mee. Het is voornamelijk gericht op cijfers en aanwezigheid.

De invloed en ontwikkeling in het gebruik van AI verschilt per mbo-sector en/of -opleiding. In een rapportage van de Samenwerkingsorganisatie Beroepsonderwijs Bedrijfsleven (SBB) worden deze verschillen uiteengezet, en wordt per sector bekeken wat de technologieën zijn die voorkomen in de verschillende sectoren en wat de output hiervan is. Daarnaast beschrijft SBB de veranderende vaardigheden die studenten nodig hebben.⁸ Zo worden digitale en computervaardigheden belangrijker. Ze benoemen wendbaarheid, ondernemendheid, creatief en kritisch denken als dragende vaardigheden, waar computational thinking, ICT-basisvaardigheden, informatievaardigheden en mediawijsheid als stuwende vaardigheden benoemd worden.

Hoger onderwijs

Op basis van een review van onderzoeksliteratuur concluderen Cardona et al. (2023) dat data uit studentvolgsystemen een belangrijke bron vormen voor technieken die het risico op studieuitval in het hoger onderwijs voorspellen. Uit de review blijkt ook dat er geen universele set variabelen is die tot de beste voorspelling leidt. Dat kan van instelling tot instelling en van studie tot studie verschillen.

⁸ Deze informatie is afkomstig uit het trendrapport dat op de website van SBB te vinden is: <https://trendrapportage.sbb.nl/trendverkenningen/kunstmatige-intelligentie/ki-algemeen/?tab=sectoroverstijgende+samenvatting&item=samenvatting>.

Fahd et al. (2021) hebben een systematische review en meta-analyses uitgevoerd op onderzoeken naar de toepassing van Machine Learning (ML, vaak een andere benaming voor algoritmen) in het hoger onderwijs. De review heeft betrekking op 89 publicaties. Het belangrijkste doel van deze reviewstudie was om meer inzicht te krijgen in de trends van de toepassing van ML in het hoger onderwijs. De meest genoemde trefwoorden in deze publicaties waren 'Machine Learning', 'Student success' en 'Classification'. In de meeste studies zijn in de modellen variabelen gebruikt die betrekking hebben op de demografische en sociaaleconomische achtergrond van de studenten en op hun pre-universitaire en universitaire studieloopbaan.

Ifenthaler en Yau (2020) voerden een reviewstudie uit met betrekking tot effecten van 'learning analytics' in het hoger onderwijs, waarbij studenten betekenisvolle terugkoppeling en ondersteuning krijgen op basis van verzamelde data over de student, het leerproces en de leeromgeving. De studie laat zien dat er een aanzienlijk aantal benaderingen met 'learning analytics' is waarin effectieve technieken worden gebruikt voor het ondersteunen van studenten. Deze technieken zouden ook studiesucces bevorderen, maar doorslaggevend bewijs daarvoor ontbreekt volgens de onderzoekers nog. Ook in deze publicatie wordt het risico op discriminatie en uitsluiting niet aangekaart. Wel wordt erop gewezen dat onderwijsinstellingen ethische en privacykwesties in aanmerking moeten nemen bij het gebruik van 'learning analytics'.

Toepassingen van algoritmen in het hoger onderwijs kwamen tijdens de meeste gesprekken met experts en stakeholders niet tot nauwelijks aan bod. In het hoger onderwijs lijkt generatieve AI, zoals ChatGPT, een grote rol te spelen. Hierbij gaat het vooral om hoe studenten deze systemen inzetten. Daarnaast wijst een expert op het gebied van technologie en privacy erop dat er juist ook tools worden ingezet die controleren of ChatGPT is gebruikt bij teksten die studenten aanleveren. In coronatijd werd veel met 'proctoring software' (of anti-spieksoftware) gewerkt, om fraude bij tentamens tegen te gaan.⁹

9 Een student van de Vrije Universiteit heeft bij het College voor de Rechten van de Mens geklaagd over de algoritmen voor gezichtsdetectie in de software 'Proctorio', waardoor zij mogelijk vanwege haar donkere huidskleur met hindernissen te maken kreeg bij het deelnemen aan tentamens tijdens de corona-periode. Uiteindelijk weerlegde de VU die suggestie echter. Meer informatie is te vinden op de website van het College voor de Rechten van de Mens: <https://www.mensenrechten.nl/actueel/nieuws/2023/10/17/student-niet-gediscrimineerd-door-tentamensoftware-proctorio-maar-vu-had-de-klacht-zorgvuldiger-moeten-behandelen>

3 Voordelen en risico's van algoritmen in het onderwijs

3.1 Inleiding

In dit hoofdstuk komen zowel de voordelen als de mogelijke risico's van algoritmen aan bod. Deze voordelen en risico's zijn thematisch gegroepeerd, omdat ze vaak ook met elkaar samenhangen.

Er zijn verschillende fasen in de constructie en het gebruik van toepassingen met algoritmen waarin rekening moet worden gehouden met risico's. Bellamy et al. (2018) onderscheiden drie fasen: 1) pre-processing, 2) in-processing, 3) post-processing. In de eerste fase kan bias ontstaan door de selectie of het gebruik van data waarmee het algoritme wordt getraind. In de tweede fase kan bias van het algoritme voor problematische uitkomsten zorgen. In de derde fase kan bias worden gecorrigeerd voordat de resultaten worden gepresenteerd. Wat precies met bias wordt bedoeld, blijkt vaak onduidelijk, zo constateren Baker en Hawn (2022). Deze auteurs geven op basis van een literatuurstudie een overzicht van verschillende soorten bias in toepassingen met algoritmen in het onderwijs. Onderzoek naar feitelijke bias in algoritmen in het onderwijs richt zich volgens de genoemde onderzoekers vooral op ras/ethniciteit, nationaliteit en geslacht. Zij concluderen onder meer dat modellen die met data van een bepaalde groep leerlingen zijn getraind, slechter lijken te presteren als zij bij een nieuwe groep leerlingen worden ingezet. Zo kunnen modellen die in het ene land zijn ontwikkeld, minder goed presteren als zij in een ander land worden ingezet. De onderzoekers wijzen er ook op dat vaak onvoldoende onderscheid wordt gemaakt binnen bepaalde diverse groepen (zoals Amerikanen van Aziatische afkomst, waarbij geen onderscheid wordt gemaakt tussen Japanse, Cambodjaanse of Indiase afkomst).

De onderzoekers wijzen erop dat er onbekende vormen van bias zijn. Het gaat dan om vormen die bepaalde groepen betreffen, zonder dat bekend is welke groepen. Zij noemen dit erg verontwaardigend en stellen: 'Without understanding which groups are at risk of algorithmic bias, even very well-intentioned attempts to resolve algorithmic bias may miss at-risk groups of students.' (Baker & Hawn, 2022, p. 1076).

Geïnterviewden geven ook aan dat veel risico's van het gebruik van algoritmen voortkomen uit het feit dat we als samenleving overspoeld zijn door de snelle opkomst en ontwikkeling van technologie en dat we dus nog moeten leren hoe daarmee om te gaan. Door de snelle ontwikkeling wordt de kennisachterstand in het onderwijs over algoritmen en AI bovendien steeds groter.

In een onderzoek van het Rathenau Instituut (2022a) dat is uitgevoerd in opdracht van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties blijkt dat men risico's bij algoritmen vooral betreft op privacy en discriminatie. Het rapport beschrijft de maatregelen die verschillende uitvoeringsinstanties nemen om de risico's bij algoritmen te verminderen. Ook doet het Rathenau Instituut aanbevelingen voor uitgangspunten voor uitvoeringsorganisaties die profilerende algoritmen inzetten of ontwikkelen.

3.2 Kansen(on)gelijkheid

Rekening houden met verschillende leerprocessen

Algoritmen kunnen helpen bij het bevorderen van kansengelijkheid op school. Dit kan bijvoorbeeld door het gebruik van adaptieve leersystemen. Een voordeel van adaptieve leersystemen is dat leerlingen hiermee beter op eigen tempo en niveau kunnen werken. Op die manier kunnen ze meer uit hun werk halen en eerlijkere kansen krijgen voor leerstofbeheersing. Eén van de geïnterviewden, die de belangen van ouders vertegenwoordigt, stelt: *‘Een goed adaptief programma kan bijdragen aan kansengelijkheid door leerlingen te helpen die geen hoogopgeleide ouders hebben.’* Dit door maatwerk mogelijk te maken en extra ondersteuning te bieden.

In het rapport *Waarden wegen* van Kennisnet wordt daarnaast benoemd dat digitale kansen niet voor alle leerlingen in funderend onderwijs en mbo gelijk zijn (Pijpers et al., 2020). Zo wordt benoemd dat het mogelijk is dat bepaalde leerlingen en studenten minder profiteren van adaptieve leersystemen. Hierbij gaat het zowel om individuele leerlingen, als om verschillen tussen scholen. Ook kan het zijn dat adaptieve software niet voor alle leerlingen het beoogde effect behaalt. Zo kan het zijn dat het voor leerlingen met dyslexie of ADHD niet even goed werkt als voor leerlingen die een dergelijke beperking niet hebben.

Programma's die teksten beoordelen van leerlingen, zoals het eerder benoemde Persuall, bevatten mogelijk een risico in het beoordelen van deze teksten. Zo kan het gebeuren dat leerlingen die Engels niet als eerste taal hebben een lagere score krijgen dan andere leerlingen, omdat het programma alleen herkent dat de woordenschat die gebruikt wordt door deze studenten niet complex genoeg is (SURF, 2022). Hierbij wordt dus geen rekening gehouden met de achtergrondinformatie over de leerling, die cruciaal is bij de beoordeling van de leerling. Een onderzoek van onderzoekers aan de Universiteit van Stanford, in de Verenigde Staten, bevestigt dat zogeheten 'GPT-detectoren' bias vertonen in het nadeel van mensen voor wie Engels niet de eerste taal is (Liang et al., 2023). Dergelijke detectoren doen op basis van algoritmen een voorspelling of een tekst door een mens is geschreven of door een 'generative pre-trained transformer' (GPT) is geproduceerd (zoals ChatGPT). De onderzoekers hebben zeven veelgebruikte GPT-detectoren ingezet om 91 door Chinese studenten in het kader van een toets Engels als vreemde taal geschreven essays en 88 door Amerikaanse studenten geschreven essays te controleren. Van de door Chinese studenten geschreven essays werd gemiddeld 61 procent onjuist gekwalificeerd als door AI geproduceerd. Hierbij bleek een vijfde van de essays door alle zeven detectoren te zijn beoordeeld als door AI gegenereerd en één van de zeven detectoren beoordeelde zelfs 98 procent van de essays incorrect. Dit terwijl de door Amerikaanse studenten geschreven teksten correct werden gekwalificeerd door de detectoren. Nadat de door Chinese studenten geschreven essays met behulp van ChatGPT waren verbeterd, daalde het percentage dat volgens de detectoren door AI was geproduceerd tot gemiddeld 12 procent. Tevens bleek na het versimpelen van de door *native speakers* geschreven essays op het gebied van idioom en grammatica dat gemiddeld 57 procent door de detectoren werd aangemerkt als door AI gegenereerd.

In een advies over het gebruik van generatieve AI-tools in het mbo, hbo en het wo schrijft de Nederlandse AI Coalitie (NL AIC) dat sommige generatieve AI-tools gebruikt kunnen worden door

studenten met een (taal)achterstand, zodat zij makkelijker bij kunnen blijven.¹⁰ Er is echter ook een scenario denkbaar waarin alleen geprivilegieerde studenten toegang hebben tot deze tools, wanneer deze bijvoorbeeld duurder worden. Daarnaast benadrukt het LAKS dat generatieve AI het lastiger maakt om de betrouwbaarheid van een (door AI) gemaakte tekst te beoordelen en bronnen te controleren. Experts van Kennisnet geven aan dat bekend is dat algoritmen die juist moeten detecteren of generatieve AI is gebruikt nog onvoldoende nauwkeurig zijn, terwijl ze wel al in de praktijk worden ingezet. Dit kan ervoor zorgen dat studenten onterecht gestraft worden omdat ze generatieve AI ingezet zouden hebben, terwijl ze dat mogelijk niet gedaan hebben of juist alleen als hulpmiddel. Bijvoorbeeld als Nederlands de tweede taal van de student is. Het risico op automatische besluitvorming is aanwezig, terwijl bekend is dat de detectiesystemen onvoldoende werken.

Representativiteit

In een onderzoek van het Centraal Planbureau (CPB) naar selectie-algoritmen in het hoger onderwijs komt daarnaast naar voren dat het mogelijk is om representatief te selecteren met behoud van effectiviteit (Hekkelman et al., 2023). Binnen dit onderzoek wordt gekeken naar vier mogelijke manieren van selecteren, te weten: een loterij, een gewogen loterij, een simpel voorspelmodel en een uitgebreid voorspelmodel. In plaats van het weglaten van gevoelige informatie, moeten de uitkomsten bijgesteld worden op een gewenste mate van representativiteit. Het weglaten van deze gevoelige informatie leidt namelijk niet altijd tot het gewenste effect. Het weglaten van bijvoorbeeld etniciteit kan ook leiden tot meer uitsluiting. Als voorbeeld benoemt het CPB dat studenten met een migratieachtergrond minder vaak doorstromen naar een specialistische artsopleiding. Een selectie-algoritme kan dit gegeven oppakken en daardoor studenten met een migratieachtergrond een lager succespercentage geven vergeleken met studenten zonder migratieachtergrond. Hierdoor worden de kansen van studenten met een migratieachtergrond juist lager om geselecteerd te worden voor een medische specialisatie.

Daarnaast schrijft het CPB dat rechtvaardigheidsoverwegingen, zoals kansengelijkheid of representativiteit, direct in de uitkomsten van een selectie-algoritme verwerkt kunnen worden. Hierbij moet opgelet worden dat er expliciete keuzes gemaakt moeten worden over welke uitkomsten nagestreefd worden.

3.3 Vooroordelen of bias

Geïnterviewden geven aan dat gevallen van discriminatie ook al voorkwamen voordat algoritmen werden gebruikt. Docenten hebben namelijk ook vooroordelen en *biases*; een goed getraind en goed toegepast algoritme kan hier mogelijk een oplossing voor bieden. De exacte uitwerking hiervan is natuurlijk algoritmespecifiek, al kunnen algemene toetsingskaders (zie paragraaf 4.3) hier een belangrijke rol in spelen. Een mogelijk voordeel van een algoritme is ook dat ze beter patronen kunnen herkennen in grote hoeveelheden data dan personen.

Een expert op het gebied van algoritmen, die zelf aan een model werkt dat de kans op studieuitval in het mbo voorspelt, zag dat docenten op basis van de houding van de student in de klas en op

¹⁰ Het advies kan worden geraadpleegd op de website van de Universiteit van Maastricht: <https://www.maastrichtuniversity.nl/file/advies-generatieve-ai-tools-het-mbo-hbo-en-wopdf>

basis van de interactie met de student bepaalde ideeën bleken te ontwikkelen over de kans op studiesucces. Het model wordt niet beïnvloed door dergelijke factoren en kan daarmee een meer objectieve inschatting geven. Door het model zodanig te ontwikkelen dat het een betere inschatting maakt dan de docent, kan tijdig worden bepaald wie extra ondersteuning nodig heeft om uitval te voorkomen. Ook de expert op het gebied van technologie en privacy stelt dat je algoritmen zou moeten gebruiken om mensen te helpen, bijvoorbeeld door vast te stellen wie extra begeleiding nodig heeft. Een eerste eis moet zijn dat bij het gebruik van algoritmen vanuit het individu wordt gedacht. Je moet het positief inzetten, bijvoorbeeld door met behulp van algoritmen te meten of mensen gelijke kansen hebben.

In een onderzoek van het Rathenau Instituut (2022b) worden discriminatie, uitsluiting en kansenongelijkheid benoemd als potentiële risico's van algoritmen. In dit onderzoek wordt specifiek gekeken naar het werken met studiedata, waarbij het mogelijk is dat een dataset waarmee een profiel wordt opgebouwd ongunstig is voor bijvoorbeeld studenten met een migratieachtergrond, waardoor deze studenten door het verwachtingsprofiel benadeeld zullen worden. Het PBLQ laat in onderzoek ook kritische geluiden horen over AI in het (mbo-) onderwijs (PBLQ, 2023). Zo benoemen ze dat men denkt dat AI onterecht 'intelligent' wordt genoemd, of dat het ongelijkheid in de hand werkt. Ook is er kritiek op de *bias* die AI kan bevatten of juist versterken. Een voorbeeld van deze *bias* is dat als je ChatGPT om de naam van een techneut vraagt, hij de naam 'Jasper' noemt, maar voor een verpleegkundige is de naam 'Elisa'. Daarnaast wordt in dit onderzoek ook de vraag gesteld wie er verantwoordelijk is voor het gebruik van AI: is dat de leverancier, het college van bestuur van een school of de docent? Ook wordt de vraag gesteld wat de rol van de docent is en wat AI doet met de sociale en emotionele interactie tussen studenten en docenten. Op deze vragen geeft het onderzoek geen antwoord.

SURF trekt de conclusie dat er in sommige AI-programma's geen rekening wordt gehouden met ondergerepresenteerde groepen, zoals etnische minderheidsgroepen of taalkundige minderheidsgroepen (SURF, 2022). In het Smarter Education with AI magazine, uitgebracht door Npuls, wordt daarnaast gesproken van een risico op nadeligheid voor minderheidsgroeperingen¹¹. Dit komt doordat AI wordt aangevuld door grootschalige data op het internet, waarbij de herkomst van de bron niet altijd goed verdeeld is. Ook Kennisnet schrijft over de aandachtspunten van AI¹². Een van de koppen die gebruikt wordt, heet *Artificial intelligence is zo goed als de trainingsdata die wordt gebruikt*. Hier legt Kennisnet uit dat AI werkt met trainingsdata (die scholen aanleveren) en dat het algoritme hierop voortborduurde. Wanneer er in de trainingsdata onjuistheden of gevoeligheden zitten, worden deze ook door het algoritme gebruikt. Dit betekent dat het gebruik van algoritmen kan leiden tot een benadeling van leerlingen, wanneer hier niet zorgvuldig mee omgegaan wordt. Om die reden heeft Kennisnet ook een publicatie¹³ over het verantwoord inzetten van AI. Hierin wordt het probleem van de slechte trainingsdata nogmaals benoemd, maar ook dat men niet blindelings de resultaten van computersystemen moet overnemen, net zoals de verschillende soorten algoritmen die er zijn. Ook het literatuuronderzoek van Velandier et al. (2023)

11 Zie ook het artikel *Ethically Responsible Application of AI in Education*, in de septembereditie van het magazine (pagina's 47 t/m 49). <https://npuls.nl/wp-content/uploads/2023/09/Smarter-education-with-AI-Npuls-september-2023.pdf>

12 Dit gebeurt in verschillende stukken op de website van Kennisnet, waaronder de volgende drie stukken: <https://www.kennisnet.nl/artikel/12989/ai-in-het-onderwijs-dit-zijn-de-belangrijkste-ethische-aandachtspunten/>, <https://www.kennisnet.nl/publicaties/artificial-intelligence/> en <https://www.kennisnet.nl/publicaties/gelijke-kansen-en-digitalisering-goede-toegang-is-niet-genoeg/>

13 Deze publicatie is eveneens op de website van Kennisnet te vinden: <https://www.kennisnet.nl/publicaties/verantwoord-inzetten-van-artificial-intelligence/>

wijst op deze risico's bij het gebruik van AI. Potentiële effecten en gevolgen worden beschreven in publicaties van Eguchi et al. (2021) en Zhou et al. (2020). De gevolgen zijn vaak ethisch gerelateerd aan bijvoorbeeld het risico op discriminatie door bevooroordeelde of niet-representatieve datasets die worden gebruikt om modellen te trainen of er is sprake van algoritmische vooringenomenheid (Arrieta et al., 2020; Mansoury et al., 2020; Mehrabi et al., 2021).

3.4 Verlichting of belasting voor onderwijspersoneel

Algoritmen kunnen worden ingezet om administratieve taken en lasten van leerlingen en medewerkers te verlichten. Automatische nakijkprogramma's besparen de leraren bijvoorbeeld vele uren nakijkwerk. Adaptieve leersystemen schotelen leerlingen gepersonaliseerde leerroutes voor, zonder dat de leraar die voor elke individuele leerling moet organiseren. Tegelijkertijd wordt door geïnterviewden aangegeven dat dit vaak oorspronkelijk het idee was van de inzet van technologie in de brede zin in het onderwijs. De laatste jaren is er echter een verandering naar meer nadruk op de mogelijke risico's van technologie, en algoritmen in het bijzonder, in het onderwijs.

Als risico van algoritmen in de onderwijspraktijk benoemen geïnterviewden veelal dat er te veel op het algoritme wordt geleund. De rol van de leraar blijft onverminderd belangrijk. Algoritmen bieden, volgens de geïnterviewden, een kans om leraren en docenten te ondersteunen, maar kunnen nooit een vervanging zijn. Eén van de geïnterviewden waarschuwt dat algoritmen zonder de menselijke maat ernaast altijd een risico vormen. Adaptieve leersystemen en toetsen zijn handig, maar de leraar moet wel kijken naar wat de leerling heeft ingevuld. Dergelijke toepassingen moeten daarom niet worden ingezet om het lerarentekort op te lossen. Deze geïnterviewde noemt nog enkele risico's van de inzet van algoritmen. Als je lang moet nadenken vóór je een antwoord geeft of als je op het toetsenbord moet zoeken naar de juiste toetsen, kan de computer daar een oordeel over je niveau aan verbinden. Dat kan ook het geval zijn als een kind een taalachterstand heeft, bijvoorbeeld door tweetalig opgroeien. Deze geïnterviewde wijst erop dat problemen bij leerlingen vluchtig kunnen zijn. Het kan dan zo zijn dat de leerling door de toepassing van algoritmen een bepaald 'label' krijgt en dat dit zo blijft terwijl het 'probleem' inmiddels niet meer bestaat. Eén van de experts wijst op het gevaar dat ontstaat als algoritmen zijn getraind op 'witte mannen'. In de Verenigde Staten bleek dat vrouwen en minderheden andere woorden gebruiken dan witte mannen. Daardoor scoren zij lager op SAT-testen (een gestandaardiseerde toelatingstest voor hogescholen en universiteiten). Daarnaast kunnen algoritmen zodanig werken dat bijvoorbeeld adhd, dyslexie of autisme ('op het spectrum') als ruis worden beschouwd door de leerprogramma's ('*Het algoritme gaat naar de gemene deler*'). De algoritmen moeten daar goed op worden getraind. Hierbij is het van belang om niet alleen aandacht te hebben voor de werking van het algoritme op populatieniveau. Juist 'aan de randen van de verdeling van de populatie', oftewel individuen met bepaalde kenmerken die niet vaak voorkomen en niet 'gemiddeld' zijn, kan een algoritme minder nauwkeurig zijn. Hier dient bij het trainen en testen van een algoritme extra aandacht aan te worden besteed. Als de achtergrond van mensen niet wordt geregistreerd (i.v.m. privacy) verdoezel je bovendien de achterliggende discriminatie.

Als leerlingen onnauwkeurig geprofileerd worden door, bijvoorbeeld, een adaptief leersysteem, kan dit leiden tot kansenongelijkheid (zoals beschreven in paragraaf 3.2). Leraren moeten in de gaten houden of het leersysteem nog aansluit bij de behoefte van de leerling. Leraren moeten daarvoor algoritmisch geletterd zijn, wat niet altijd zo is. In hoeverre leraren algoritmisch geletterd moeten zijn, is ook vaak punt van discussie. Kennisnet geeft hierbij aan dat gebruikers van

algoritmen geholpen zijn bij een heldere, simpele en concrete uitleg. Dit kan bestaan uit een goede bijsluiter bij het gebruikte algoritme. In deze bijsluiter kan kennis over het algoritme worden overgedragen, inclusief overwegingen die zijn gemaakt. Goede betrouwbare informatie op het juiste detailniveau is dus belangrijk. Daarnaast is het van belang dat nog gedetailleerdere vragen ook worden beantwoord, door bijvoorbeeld de ontwikkelaar van het product.

Ontwikkelaars geven aan dat adaptiviteit ervoor zorgt dat leerlingen vragen krijgen die beter bij de vaardigheden van de betreffende leerling aansluiten. Hierdoor hoeft een leerling ook minder vragen te beantwoorden om het niveau vast te stellen. Daartegenover staat dat LAKS aangeeft dat ze signalen ontvangen dat adaptieve leersystemen juist demotiverend werken als vragen fout worden beantwoord. Het voelt voor leerlingen alsof ze dan extra oefeningen moeten uitvoeren. Daarnaast ontbreekt de terugkoppeling vaak hoe een adaptief leersysteem tot een bepaalde route komt.

3.5 Privacy

De inzet van 'learning analytics' zonder duidelijk doel en heldere kaders vergroot de kans op onrechtmatige of onevenredige inbreuken aanzienlijk. Data en profielen gebruiken zonder duidelijkheid over wat de data wel en niet betekenen en welke conclusies eruit getrokken mogen worden, vergroot het risico op onrechtvaardige behandeling van studenten, schrijft de Autoriteit Persoonsgegevens (2023a). Ook blijkt uit de interviews met vertegenwoordigers van leerlingen en studenten (LAKS en JOB), dat in de praktijk leerlingen/studenten vaak geen weet hebben van het gebruik van hun data voor dergelijke algoritmen.

Privacykwesties spelen ook in overvloed in het complexe domein van eigendom en aggregatie van gegevens (Kearns & Roth, 2020). Aangezien de mechanismen achter datagestuurde praktijken, zoals het verzamelen en analyseren van gegevens, vaak verborgen zijn voor gebruikers en de analysemethoden vaak zeer moeilijk uit te leggen zijn (Arrieta et al., 2020), zijn gebruikers zich vaak niet bewust van de mogelijke impact die deze praktijken hebben op hun leven en hun keuzes (Pangrazio & Selwyn, 2021).

4 Huidig en wenselijk beleid

4.1 Inleiding

In de fase van deskresearch is nagegaan of verschillende betrokken organisaties en instellingen beleid voeren op het gebied van algoritmen in het onderwijs. Ook in de vraaggesprekken is daar aandacht aan besteed. In die gesprekken is ook de vraag gesteld wat het beleid op dat gebied zou moeten inhouden en wat onderwijsorganisaties en -instellingen zelf zouden kunnen en moeten doen.

4.2 Beleid van sectororganisaties en andere betrokken instanties

De sectororganisaties in het primair, voortgezet en middelbaar beroepsonderwijs

Navraag bij de verschillende sectororganisaties in het onderwijs (PO-Raad, VO-raad, MBO Raad) wijst uit dat er bij die organisaties weinig tot geen beleid is op het gebied van het gebruik van algoritmen in het onderwijs. Een deel van de organisaties geeft aan hier wel mee bezig te zijn, een ander deel heeft nog geen stappen in gezet. Daarnaast zijn de sectororganisaties wel vaak op de hoogte van andere partijen die bezig zijn met het onderwerp, zoals de Inspectie van het Onderwijs of SURF. Dit zijn veelal partijen waar zij vaak mee samenwerken of vaker mee in contact staan.

Vertegenwoordigers van het hoger onderwijs

Ook uit navraag bij de Vereniging Hogescholen en bij Universiteiten van Nederland (UNL) bleek dat daar weinig beleid is op het gebied van algoritmen in het onderwijs. Wel bleken die partijen meer kennis te hebben van algoritmen (of digitale toepassingen) in hun omgeving. De Vereniging Hogescholen weet bijvoorbeeld dat er verschillende individuele instellingen zijn die wel bezig zijn beleid te vormen op dit terrein. Bij UNL is er weliswaar geen beleid op algoritmen in het onderwijs, maar UNL is wel bezig met digitale toepassingen en AI in het onderwijs. Zo wordt er gewerkt aan het opstellen van een overzicht van publieke waarden en academische soevereiniteit waar het gaat om AI in het onderwijs. Ook is UNL bezig met het verankeren van deze publieke waarde in wet- en regelgeving door middel van een lobby. UNL onderscheidt drie waarden die zij belangrijk noemen voor het onderwijs. Dit zijn rechtvaardigheid, menselijkheid en autonomie¹⁴. Dit kan door de verschillende universiteiten verder worden ingevuld in hun strategie en (investerings)beslissingen. Daarnaast zijn er verschillende universiteiten die eigen richtlijnen hebben opgesteld voor het werken met generatieve AI, zoals ChatGPT.

De UNL deelt zeven waarden voor *open research information*, opgesteld door de werkgroep publieke waarden¹⁵. Deze waarden zijn opgesteld om ervoor te zorgen dat de academische gemeenschap de wetenschappelijke waarden voor metadata-infrastructuren handhaaft. Een van de zeven waarden is openheid van algoritmen. 'Black-box' algoritmen belemmeren transparante,

14 Deze waarden worden besproken in een advies op de website van Universiteiten van Nederland: <https://www.universiteitenvannederland.nl/files/publications/Advies%20werkgroep%20publieke%20waarden%20onderwijs.pdf>

15 Deze waarden zijn op de volgende website uitgewerkt: <https://zenodo.org/records/6074944>

eerlijke besluitvorming, bijvoorbeeld bij keuzes op het gebied van wetenschappelijke evaluatie en erkenning. Deze waarden zijn vooral gericht op wetenschappelijk onderzoek dat uitgevoerd wordt op de universiteiten, maar raakt ook het onderwijs.

Een van de experts wijst erop dat er een referentiekader privacy en ethiek is voor het hoger onderwijs en dat er een wordt ontwikkeld voor het mbo. Ook in het gesprek met JOB wordt aangegeven dat in het mbo op initiatief van de sector zelf een referentiekader wordt ontwikkeld waarin wordt vastgelegd hoe om te gaan met studiedata.

Daarnaast was er tussen 2019 en 2022 het Versnellingsplan Onderwijsinnovatie met ICT, dat een referentiekader heeft geschreven voor het veilig en betrouwbaar benutten van studiedata (Versnellingsplan, 2021). Binnen dit referentiekader worden vier ethische uitgangspunten en vier juridische privacy-uitgangspunten gedeeld over het gebruik van studiedata. Hieronder valt ook het ruimte houden voor de menselijke maat, ook wanneer er gebruik wordt gemaakt van geautomatiseerde processen, alsook aanspreekbaar en transparant zijn over het gebruik van studiedata.

Vertegenwoordigers van leerlingen en studenten

In de vraaggesprekken geven vertegenwoordigers van de scholieren en studenten (LAKS, ISO en JOB) aan dat er bij de sectororganisaties weinig tot geen beleid is voor de ontwikkeling en inzet van algoritmen in het onderwijs. Tegelijkertijd benadrukken ze wel het belang hiervan. Dit beleid zou zich moeten richten op verschillende aspecten, zoals het voorkomen van discriminatie en het vergroten van de aandacht en bewustwording van de risico's van het gebruik van algoritmen. Ook geven zij aan dat individuele scholen weinig beleid hebben op dit terrein en transparanter moeten zijn. Vanuit Ouders & Onderwijs wordt ook opgemerkt dat ouders slechte informatie vanuit school krijgen, bijvoorbeeld als er een schoolcomputer thuis wordt gebruikt met een toezichtprogramma (bij het maken van toetsen). Wat volgens de vertegenwoordiger van Ouders & Onderwijs steeds tegenvalt, is dat gebruikers weinig kennis hebben van hoe een programma werkt. Dit sluit ook aan bij de in paragraaf 2.2 vermelde conclusie uit de MYRA-monitor dat schoolbestuurders, schoolleiders en leraren zelf aangeven weinig kennis te hebben van AI (Karssen et al, 2023a, 2023b).

Andere organisaties

De ICT-coöperatie SIVON behartigt vanuit publieke waarden de belangen van haar leden, scholen in het primair en voortgezet onderwijs. SURF vervult deze rol voor onderwijsinstellingen zoals universiteiten, hogescholen, mbo-instellingen, umc's en onderzoeksinstellingen. Daarnaast ondersteunt Kennisnet scholen bij een professionele inzet van ICT, adviseert de PO-Raad, de VO-raad, de MBO Raad en MBO Digitaal en ondersteunt grootschalige bestuurlijke initiatieven die de inzet van ICT pogen te professionaliseren. Ook publiceert Kennisnet artikelen over het (verantwoordelijk) gebruik van AI in het onderwijs. Kennisnet schrijft daarnaast achtergrondartikelen voor professionals in het onderwijsveld over AI en algoritmen.

Beleid van de overheid

Na een lange aanloop is de 'Artificial Intelligence Act' (AI Act) van de Europese Unie goedgekeurd door de lidstaten. Daarmee wordt onder meer beoogd dat AI-systemen veilig zijn en de rechten van burgers niet schenden. De AI-verordening verbiedt AI in het onderwijs te gebruiken voor het herkennen van emotie. Daarnaast worden een aantal toepassingen in het onderwijs als hoogrisico beschouwd, zoals systemen die worden gebruikt om toelating tot onderwijsinstellingen te bepalen, om leerresultaten te evalueren, ook wanneer die resultaten worden gebruikt voor het

sturen van het leerproces, en AI die gebruikt wordt om ongeoorloofd gedrag te monitoren tijdens het maken van examens.

Ook is er een landelijk algoritmeregister. Daarin moet aangegeven worden op basis van welke informatie uitkomsten worden gegenereerd. Daarmee is een poging tot transparantie gedaan, zo geeft een van de geïnterviewde experts aan. Ook is er een impact assessment voor mensenrechten bij de inzet van algoritmen (IAMA) (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2021). Dit is een instrument voor discussie en besluitvorming voor overheidsorganen bij de afweging om wel of niet een algoritmische toepassing te gaan ontwikkelen. Het gebruik van dit instrument moet een interdisciplinaire dialoog mogelijk maken door degenen die verantwoordelijk zijn voor de ontwikkeling en/of inzet van een algoritmisch systeem.

Uit een verkenning naar een Nederlands Onderwijslab voor AI (Bureau Common Ground & it's public, 2020) bleek dat er nog weinig informatie beschikbaar is over het gebruik van AI in de klas. Dit was een van de redenen om het Nationaal Onderwijslab AI op te zetten. Het NOLAI richt zich op het basisonderwijs, voortgezet en speciaal onderwijs. Het is verbonden aan de Radboud Universiteit en wordt gefinancierd door het Nationaal Groeifonds. Binnen het NOLAI worden algoritmen (intelligente technologieën) ontwikkeld in samenwerking tussen wetenschappers, scholen en mensen uit het bedrijfsleven.

De Inspectie van het Onderwijs geeft in het vraaggesprek aan dat toezicht op AI (waaronder algoritmen) nog in de kinderschoenen staat. Verschillende toezichthouders (Inspectie van het Onderwijs, Autoriteit Persoonsgegevens en de Auditdienst Rijk) overleggen momenteel samen hoe om te gaan met toezicht op algoritmen. De Inspectie van het Onderwijs geeft aan voornamelijk gericht te zijn op toezicht op de onderwijskwaliteit. De Autoriteit Persoonsgegevens heeft wellicht meer technische kennis die nodig is voor toezicht op algoritmen in het onderwijs.

CitoLab en Bureau ICE geven in de vraaggesprekken aan dat er beleid is voor systemen met een wettelijk verplichtend karakter, zoals de doorstroomtoets. Sommige doorstroomtoetsen zijn adaptief, of maken gebruik van beslisregels. Dit worden algoritmen genoemd in het beleid van de overheid. Dit beleid zorgt ervoor dat deze algoritmen goed worden gecontroleerd en dat algoritmen op een goede manier worden ingezet¹⁶. De Regeling beoordelingskader doorstroomtoets PO beschrijft waaraan deze algoritmen moeten voldoen. Er moet kunnen worden verantwoord dat ankeropgaven verantwoord voortkomen. De afname van de doorstroomtoets moet aan eisen voldoen ten aanzien van het gezamenlijk anker, op een manier zoals beschreven in de hiervoor genoemde Regeling.

4.3 Welk beleid is wenselijk op het gebied van algoritmen in het onderwijs?

In de verschillende onderzoeksfasen is aandacht besteed aan de vraag wat er op beleidsgebied zou moeten gebeuren om de risico's van het gebruik van algoritmen in het onderwijs te vermijden of te ondervangen. Daarbij gaat het niet alleen om beleid van de overheid, maar ook om beleid van de onderwijssectoren en van individuele scholen en onderwijsinstellingen. Uiteraard zijn er

¹⁶ Meer informatie over deze beoordeling is te vinden in de Regeling beoordelingskader doorstroomtoets PO: <https://wetten.overheid.nl/BWBR0047614/2024-03-01>

ook voorzorgsmaatregelen die ontwikkelaars en aanbieders van toepassingen met algoritmen kunnen nemen om risico's te ondervangen of te verkleinen.

Toetsingskaders

Er bestaan verschillende toetsingskaders voor algoritmen. Zo heeft de Algemene Rekenkamer het Toetsingskader Algoritmes ontwikkeld.¹⁷ Dit toetsingskader is een praktisch instrument dat gebruikt kan worden om te toetsen of algoritmen aan bepaalde kwaliteitscriteria voldoen en of er genoeg bekend is over de mogelijke risico's. Een ander voorbeeld is 'FATE'. Dat staat voor 'Fairness, Accountability, Transparency, Ethics' (Memarian & Doleck, 2023). De stakeholders op het gebied van de ontwikkeling van algoritmen zouden in FATE moeten worden onderwezen. Bogina et al. (2022) onderscheiden drie typen stakeholders: ontwikkelaars, degenen die zich met regelgeving bezighouden en gebruikers. Binnen die laatste groep wordt onderscheid gemaakt tussen de systeemeigenaars en de eindgebruikers. Memarian en Doleck (2023) voerden een literatuurstudie uit naar definities van FATE in het hoger onderwijs en onderzochten tevens of de definities descriptief of technisch van aard waren. Fairness bleek verreweg het meest besproken thema van de vier onderscheiden begrippen. Transparency en Ethics kwamen het minst aan bod.

Wenselijk beleid van de overheid

Zoals aangegeven, wordt met de AI Act van de Europese Unie beoogd om het gebruik van AI – en daarmee ook van algoritmen in het onderwijs – veiliger te maken. Het ministerie van OCW geeft aan dat de AI Act goede handvatten biedt voor het reguleren van algoritmen in het onderwijs. Hier ligt echter wel een belangrijke taak voor de komende tijd om een verdiepingsslag op basis van de AI Act te maken, waarbij de vertaling naar de onderwijspraktijk gemaakt dient te worden.

Een expert op het gebied van technologie en privacy geeft in het vraaggesprek echter aan dat de Europese normen te algemeen zijn. Volgens die normen moet je jezelf reguleren door een *impact assessment* te doen. Scholen hebben niet de expertise om dat te doen en zij hebben bovendien te weinig invloed op de leveranciers van toepassingen. Deze expert stelt dat je algoritmen 'tot op het bot' moet testen, op grote schaal, in een 'sandbox-achtige situatie'. Zij noemt 'SLM Rijk' als goed voorbeeld. SLM Rijk is de centrale onderhandelaar voor producten en diensten van Microsoft, Google en Amazon Web Services voor de rijksoverheid.¹⁸ SLM staat voor Strategisch Leveranciersmanagement. Zij voeren een Data Protection Impact Assessment (DPIA), Transfer Impact Assessment (TIA) en Human Rights Assessment uit en kunnen eisen stellen aan leveranciers. Zij stellen de voorwaarden vast en publiceren zelf DPIA's. Volgens deze expert hebben DPIA's meer effect gehad op de wereldwijde privacy dan welke wettelijke maatregel ook. De overheid of een organisatie als SURF zouden Strategisch Leveranciersmanagement moeten uitvoeren op producten die voor het onderwijs bestemd zijn. Als een product dan wordt goedgekeurd, weten scholen dat ze het kunnen gebruiken. Je kunt risico's verminderen door heel diepgaande algoritmen open te stellen voor inspecties door academici. Bij inkoop moet je daarop staan.

17 Informatie over het Toetsingskader algoritmes is te vinden op de website van de Rekenkamer: <https://www.rekenkamer.nl/onderwerpen/algoritmes/algoritmes-toetsingskader>

18 Er is een speciale website waar DPIA's, technische rapporten en handleidingen m.b.t. het kiezen van de juiste instellingen kunnen worden gedownload: <https://slmmicrosoftrijk.nl/>

Het Rathenau Instituut doet in haar onderzoek aanbevelingen voor onderzoek naar leeropbrengsten (Rathenau Instituut, 2022b). Hierbij wordt aanbevolen om te overwegen kwaliteitskeurmerken in te voeren voor digitale leermiddelen. Deze keurmerken moeten onderwijsinstellingen helpen wegwijs te worden in digitale onderwijsmiddelen.

De Verenigde Naties (VN) erkent dat AI en algoritmen kunnen leiden tot discriminatie (Unesco, 2021). Zij spreken over het opzetten van monitorende agentschappen om te controleren op andere discriminatie bij AI. Daarnaast benoemen ze dat de overheid via antidiscriminatiewetgeving erop moet toezien dat instellingen die analyses uitvoeren op data van leerlingen en studenten, uitgaan van algoritmen die vrij zijn van vertekening en alle maatregelen hebben genomen om de veiligheid van hun data te borgen. Daarnaast schrijft ook de Onderwijsraad (2022) dat de overheid een taak heeft om de privacy en dataveiligheid te waarborgen. Ook moet de overheid leerdoelen formuleren over de digitale geletterdheid van leerlingen, waar het gebruik van digitale leermiddelen en AI onderdeel van zijn.

Wenselijk beleid van en voor onderwijssectoren en -instellingen

De Onderwijsraad wijst erop dat het de taak is van schoolleiders en bestuurders om ervoor te zorgen dat er geen ongelijke kansen of discriminatie ontstaan door het gebruik van intelligente technologie (Onderwijsraad, 2022). Binnen onderwijsinstellingen moet geborgd worden dat er geen voordeel voor een deel van de leerlingen ontstaat. Twee experts wijzen er in de vraaggesprekken op dat binnen de Europese 'data protection act' scholen onder 'high risk' vallen. Er moet dus aan strenge eisen worden voldaan. De eisen aan het systeem moeten in contracten worden vastgelegd. Eén van de experts geeft ook aan dat je voor de gebruiker duidelijk moet maken welke instellingen nodig zijn in het programma, om te voorkomen dat de toepassing alsnog risicovolle acties uitvoert. De expert die werkzaam is in het mbo stelt dat het handig zou zijn als de MBO Raad of een ander overkoepelend orgaan met een opzet komt waarin staat waar onderwijsinstellingen mee moeten werken.

Eén van de experts is van mening dat je het onderwijs een openingsperspectief moet bieden, vanwege het gebrek aan kennis over algoritmen. Via een filmpje zou je onderwijsgevendend bewust moeten maken van de risico's. Dat moet specifiek gericht zijn op het product dat ze zelf gebruiken en het moet een voorbeeld geven van hoe je dat product positief kunt inzetten. Een expert die zelf een model ontwikkelt dat studieuitval kan voorspellen, wijst erop dat het om onderwijsgevendend bewust te maken van de mogelijkheden en risico's van algoritmen belangrijk is dat ze digitaal vaardig zijn en weten wat een algoritme inhoudt en wat de gevaren kunnen zijn. Ook bij de lerarenopleidingen zou daar aandacht voor moeten zijn.

Een aantal betrokkenen benadrukt in de vraaggesprekken het belang van beleid en controle, aangezien het onwenselijk is om commerciële partijen te veel vrijheid te bieden. Beleid hoeft echter niet alleen vanuit de overheid te komen. De eigen verantwoordelijkheid van scholen/instellingen wordt ook benadrukt door stakeholders en experts. Scholen/instellingen zijn bijvoorbeeld een belangrijke vrager op de markt van digitale leermiddelen. Tegelijkertijd wordt erkend dat een individuele school geen positie heeft ten opzichte van grote aanbieders en weinig eisen kan stellen aan deze aanbieders.

Daarnaast benadrukken tal van de geïnterviewden ook het belang van transparantie. Wat algoritmen precies doen, blijft vaak verborgen. Er zijn commerciële belangen mee gemoeid ('*Het is moeilijk om toezicht te houden als iets niet zichtbaar is*'). Het is op dit moment vaak onduidelijk

welke gegevens voor welk algoritme worden gebruikt en hoe een algoritme tot stand komt. Hier zou meer transparantie over moeten zijn voor de gebruikers. Een van de stakeholders stelt dat het goed zou zijn als educatieve software open source zou zijn. Dan is transparantie gewaarborgd. Als dat niet haalbaar is, dan is het in ieder geval belangrijk dat scholen zelf eigenaar van de verzamelde data zijn en dat de data niet door partijen met eigen belangen worden opgeslagen.

Een expert die in het mbo werkzaam is, wijst erop dat op scholen naar de ethische kant zou moeten worden gekeken en er zouden met het hele team afspraken moeten worden gemaakt over spelregels. Verder moeten studenten de mogelijkheid hebben om 'opt out' te kiezen als er algoritmen worden ingezet. Scholen of onderwijsinstellingen moeten aangeven dat die mogelijkheid er is. Ook in de wetenschappelijke literatuur wordt gewezen op de rechten van leerlingen en studenten. Zij verschaffen de data die worden verzameld, dus zouden zij ook het recht op eigenaarschap moeten hebben en invloed kunnen uitoefenen op wat er met hun data gebeurt, zodat zij er bij het leren profijt van hebben (Holmes et al., 2021).

Kennisnet heeft een Ethiekkompas, waar scholen of scholengroepen een vraag kunnen stellen of een casus kunnen bekijken die te maken heeft met digitale toepassingen in het onderwijs. Op deze manier kunnen scholen bekijken of de keuze voor digitale toepassingen geen onbedoelde negatieve effecten heeft.

5 Conclusies

5.1 Inleiding

Tot besluit worden in dit hoofdstuk de belangrijkste conclusies van het onderzoek samengevat, worden de onderzoeksvragen beantwoord en worden aanbevelingen gegeven.

5.2 Conclusies

Toepassingen van algoritmen in het onderwijs

Algoritmen worden in tal van toepassingen en in alle onderwijssectoren gebruikt, het meest in het primair en hoger onderwijs. In het middelbaar beroepsonderwijs en het hoger onderwijs wordt onder meer gebruik gemaakt van 'learning analytics', om de studievoortgang te volgen en de kwaliteit van het onderwijs te monitoren. In het primair en voortgezet onderwijs worden adaptieve leersystemen gebruikt, evenals automatische nakijkprogramma's, dashboards, simulaties, leerlingvolgsystemen en adaptieve toetsen. In het mbo en hoger onderwijs wordt gewerkt aan modellen die vroegtijdig studieuitval en de behoefte aan extra ondersteuning kunnen voorspellen. Ook zijn er niet op de inhoud gerichte toepassingen met algoritmen. Een voorbeeld zijn plaatsingssystemen bij de overgang van primair naar voortgezet onderwijs. Een ander voorbeeld zijn toepassingen die zijn gericht op surveillance, zoals software die kan worden ingezet om met behulp van generatieve AI geproduceerde teksten te herkennen en 'proctoring software' om fraude bij tentamens te detecteren. De Inspectie van het Onderwijs gebruikt algoritmen om risico-inschattingen van scholen te maken. Zelflerende algoritmen worden (nog) niet in de onderwijspraktijk gebruikt. De algoritmen die momenteel in gebruik zijn, werken op basis van geautomatiseerde beslisregels die geprogrammeerd zijn, al dan niet in combinatie met data waarmee het algoritme is getraind.

Voordelen

Een mogelijk voordeel van algoritmen is dat deze ertoe kunnen leiden dat taken efficiënter worden uitgevoerd dan zonder algoritmen het geval is. Algoritmen zijn in staat snel patronen te herkennen in grote hoeveelheden data. Algoritmen kunnen administratieve taken deels overnemen en lasten van de leraar verlichten. Een goed werkend algoritme kan worden ingezet om te voorspellen welke studenten of leerlingen extra ondersteuning nodig hebben. Dankzij adaptieve leersystemen – mits zonder bias – kunnen leerlingen beter in eigen tempo en op eigen niveau werken en degenen die dat nodig hebben, kunnen extra uitleg of ondersteuning krijgen. Dit vanuit de digitale toepassing zelf of doordat de leraar een signaal krijgt dat extra ondersteuning nodig is. Bij een goed werkend algoritme kunnen leerlingen meer uit hun werk halen en eerlijkere kansen krijgen. Een goed algoritme zou kunnen bijdragen aan kansengelijkheid van leerlingen. Een ander voordeel van algoritmen – mits goed getraind en goed toegepast – is dat deze niet door eventuele vooroordelen van de docent ten aanzien van studenten of leerlingen worden beïnvloed. Bij de juiste inzet kunnen algoritmen bijdragen aan het tegengaan van vooroordelen van docenten.

Risico's

Risico's die bij algoritmen worden gezien, hebben vooral betrekking op inbreuken op de privacy, discriminatie, uitsluiting en kansenongelijkheid. Dergelijke risico's kunnen het gevolg zijn van

tekortkomingen in drie fasen van het proces van ontwikkeling en toepassing. Bias kan ontstaan door de selectie of het gebruik van data waarmee het algoritme wordt getraind en bias kan ontstaan door de manier waarop het algoritme is geprogrammeerd. Na de inzet van het algoritme kunnen risico's ontstaan als de uitkomsten niet op bias worden gecontroleerd. Verschillende betrokkenen wijzen er daarom op dat de inzet van adaptieve leersystemen niet mag inhouden dat de leraar niet meer toeziet op wat de leerlingen doen. Een ander probleem van algoritmen is dat er vaak een gebrek aan transparantie is. Daardoor is onduidelijk hoe ze werken en kunnen eventuele gevallen van discriminatie of uitsluiting verborgen blijven voor de gebruiker. Ook is vaak onduidelijk wat er aan data wordt opgeslagen, waar dat gebeurt en of daar vervolgens nog meer mee gebeurt. Overigens wordt er ook op gewezen dat eventuele correctie door een algoritme voor kansenongelijkheid niet mogelijk is als bepaalde als gevoelig beschouwde achtergrondinformatie wordt weggelaten. Wat als een belangrijk bezwaar wordt genoemd, is dat algoritmen zich richten naar de gemene deler, als hier niet voor wordt gecorrigeerd. Dat kan risicovol zijn voor bepaalde groepen, zoals leerlingen uit minderheidsgroepen en leerlingen met beperkingen.

Beleid

Uit een rondgang bij betrokken organisaties en instanties blijkt dat er weinig beleid is op het gebied van algoritmen in het onderwijs. Sommige partijen zijn daar wel mee bezig. Met de Europese AI Act wordt gepoogd de risico's van artificiële intelligentie in te dammen, maar experts wijzen erop dat die regels te algemeen zijn en nog om een specifieke invulling vragen. Er is een algoritmeregister van de Nederlandse overheid, waarmee wordt getracht tot meer transparantie te komen over toepassingen met algoritmen en hoe deze te werk gaan. Er is een referentiekader privacy en ethiek voor het hoger onderwijs en er wordt er een ontwikkeld voor het mbo. Vertegenwoordigers van scholieren en studenten zien ook dat er weinig beleid op het gebied van algoritmen is bij sectororganisaties, terwijl dat wel belangrijk wordt geacht. Vertegenwoordigers van ouders merken dat ouders slecht door scholen worden geïnformeerd over toepassingen waarin algoritmen aan het werk zijn.

Wenselijk beleid

Er zijn verschillende toetsingskaders voor algoritmes. In publicaties wordt, bijvoorbeeld, gewezen op het belang van 'Fairness, Accountability, Transparency, and Ethics' (FATE), zowel voor ontwikkelaars als voor degenen die verantwoordelijk zijn voor regelgeving, de eigenaars van de systemen en de gebruikers. Algoritmen moeten op grote schaal worden getest en er moeten afspraken worden gemaakt met leveranciers. Idealiter gebeurt dat niet op schoolniveau, maar op landelijk niveau. Overheidsorganisaties en/of vertegenwoordigers van de onderwijssectoren zouden assessments (waaronder DPIA's) kunnen uitvoeren en eisen stellen aan leveranciers. Als aan alle voorwaarden is voldaan, zou een keurmerk kunnen worden afgegeven. Als toepassingen een keurmerk krijgen, is voor onderwijsinstellingen duidelijk dat deze betrouwbaar zijn. Daarbij is ook van belang wie de eigenaar van de verzamelde data is. Als met algoritmen wordt gewerkt, moeten ouders, leerlingen en studenten daarover worden geïnformeerd, evenals over welke data worden opgeslagen. Ook moet hun toestemming worden vastgelegd. Sectororganisaties zouden scholen en onderwijsinstellingen kunnen ondersteunen door vast te leggen aan welke voorwaarden moet worden voldaan.

Verder wordt gewezen op het belang dat onderwijsgevendende voldoende getraind worden in het werken met toepassingen waarin algoritmen worden gebruikt en dat zij ook de risico's daarvan moeten leren onderkennen. Daarin zouden ook studenten aan de lerarenopleidingen moeten worden onderwezen.

5.3 Beantwoording van de onderzoeksvragen

Zoals in paragraaf 1.3 is aangegeven, is voor dit onderzoek één hoofdvraag gesteld. Deze luidt als volgt: *Welke risico's op discriminatie en uitsluiting (en kansen ter voorkoming hiervan) doen zich voor bij de inzet van algoritmen in het onderwijs in Nederland of kunnen zich in de toekomst voor gaan doen?*

Vervolgens zijn op basis van de hoofdvraag zeven deelvragen geformuleerd. Deze zeven deelvragen worden hieronder beantwoord. Daarmee wordt tevens het antwoord gegeven op de hoofdvraag.

1) Hoe worden algoritmen binnen het onderwijs ingezet of wordt de inzet daarvan overwogen?

Algoritmen worden op verschillende manieren ingezet. Er zijn toepassingen met algoritmen die op de onderwijsinhoud zijn gericht, zoals adaptieve leersystemen en automatische nakijksystemen. Er zijn toepassingen die op het volgen van leerlingen en hun prestaties zijn gericht, zoals adaptieve toetsen en leerlingvolgsystemen. Er zijn toepassingen die de kans op studie-uitval en de behoefte aan extra ondersteuning inschatten. Er zijn toepassingen voor plaatsing van leerlingen op een school en er zijn toepassingen die zijn gericht op het detecteren van fraude.

2) Welke risico's op discriminatie en/of uitsluiting door de inzet van algoritmen zijn gedetecteerd in het onderwijs in Nederland?

Er worden risico's gesignaleerd, zowel in de internationale wetenschappelijke literatuur als door experts en andere betrokkenen in Nederland. Daarbij wordt bijvoorbeeld verwezen naar leerlingen uit minderheidsgroepen en leerlingen met beperkingen die zouden kunnen worden benadeeld door uitkomsten van algoritmen en/of beslissingen die op basis daarvan worden genomen. In het onderzoek zijn geen concrete voorbeelden van discriminatie en/of uitsluiting door algoritmen in Nederland gevonden.

3) Welke risico's op discriminatie en/of uitsluiting door de inzet van algoritmen zijn niet gedetecteerd, maar kunnen verondersteld worden?

Er zijn geen concrete voorbeelden van discriminatie en/of uitsluiting door het gebruik van algoritmen gevonden in Nederland. De gesignaleerde risico's worden dus verondersteld. Een onderzoek in de Verenigde Staten liet zien dat Engelstalige teksten die zijn geschreven door personen voor wie Engels niet de eerste taal is, door detectoren op basis van algoritmen in meer dan de helft van de gevallen werden gekwalificeerd als door AI geproduceerd. Op basis hiervan kan worden verondersteld dat leerlingen en studenten voor wie Nederlands niet de eerste taal is, ten onrechte van fraude zouden kunnen worden verdacht bij de inzet van dergelijke detectoren om door hen in het Nederlands geschreven teksten te controleren.

4) Welke risico's van de inzet van algoritmen zouden in de toekomst kunnen toenemen?

Door de snelle ontwikkelingen in de technologie, de brede inzet daarvan in het onderwijs, onvoldoende kennis bij beleidsmakers en gebruikers in het onderwijs en gebrek aan transparantie over hoe algoritmen functioneren, ligt het voor de hand dat alle risico's – bij ongewijzigd beleid – in de toekomst toenemen.

5) In welke fase(n) en/of bij welke type(n) van onderwijs spelen deze risico's met name?

Deze risico's spelen bij alle fasen en onderwijssectoren. Uit het onderzoek blijkt dat algoritmen met name in het primair en hoger onderwijs worden toegepast, dus daar zouden de risico's sterker kunnen spelen dan in het voortgezet en speciaal onderwijs en het middelbaar beroepsonderwijs.

6) Wat is de stand van zaken van het beleid van het Ministerie van Onderwijs, de Inspectie van het Onderwijs en scholen op het gebied van het voorkomen van discriminatie en/of uitsluiting door de inzet van algoritmen in het onderwijs?

De Inspectie van het Onderwijs richt zich primair op het toezicht op de onderwijskwaliteit. Toezicht op algoritmen is dan pas aan de orde als deze de onderwijskwaliteit beïnvloeden. De Autoriteit Persoonsgegevens en de Rijksdienst Digitale Infrastructuur hebben een werkgroep opgezet die verkent welke toezichthouders toezicht moeten houden op specifieke toepassingen van AI. De Inspectie wordt hier ook bij betrokken omdat het onderwijs wordt genoemd in de Europese AI Act. Scholen en onderwijsinstellingen hebben weinig beleid op het gebied van de inzet van algoritmen in het onderwijs.

7) Wat kunnen het Ministerie van Onderwijs, de Inspectie van het Onderwijs en onderwijsinstellingen (meer, anders) doen om de (verkende risico's op) discriminatie en/of uitsluiting door de inzet van algoritmen te voorkomen?

Er moeten strenge eisen worden gesteld aan toepassingen. Hierbij heeft de overheid een belangrijke functie. Softwareontwikkelaars zijn ervoor verantwoordelijk dat aan die eisen wordt voldaan. Toepassingen met algoritmen moeten zeer uitgebreid worden getest op mogelijke negatieve uitwerking voor bepaalde groepen. Dat moet allereerst door de softwareontwikkelaar worden gedaan. Vervolgens moet op landelijk niveau worden getest of systemen aan de voorwaarden voldoen. Als dat het geval is, zou er een keurmerk voor kunnen worden afgegeven. Scholen missen de expertise om uitgebreid te testen en hebben onvoldoende invloed bij de leveranciers van toepassingen. Onderwijsgeevenden zouden beter moeten worden geschoold in het werken met technologie in het onderwijs. Ook zouden zij meer bewust moeten worden gemaakt van potentiële risico's van algoritmen en manieren om die te ondervangen.

5.4 Aanbevelingen

Op basis van dit onderzoek formuleren wij de volgende aanbevelingen:

- stimuleer transparantie over algoritmen in toepassingen voor het onderwijs;
- stel voorwaarden op waaraan toepassingen voor het onderwijs waarin algoritmen werkzaam zijn moeten voldoen;
- zorg ervoor dat toepassingen met algoritmen uitgebreid worden getest op onwenselijke acties en uitkomsten voor specifieke groepen;
- maak afspraken met leveranciers van toepassingen die data verzamelen over het eigendom van die data;
- voer een landelijk keurmerk in voor toepassingen die voldoen aan de eisen;
- zorg ervoor dat de lerarenopleidingen toekomstige onderwijsgeevenden beter voorbereiden op het werken met toepassingen voor het onderwijs waarin algoritmen zijn ingebouwd;
- zorg ervoor dat onderwijsgeevenden beter worden geschoold in het werken met technologie in het onderwijs en met toepassingen waarin algoritmen zijn ingebouwd;

- zorg ervoor dat onderwijsgeevenden meer bewust worden gemaakt van de risico's van algoritmen en manieren waarop zij die kunnen verkleinen of ondervangen;
- ga in een vervolgonderzoek na wat de feitelijke positieve en negatieve effecten in de onderwijspraktijk zijn van het gebruik van algoritmen.

Literatuurlijst

- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
- Autoriteit Persoonsgegevens (2023a). *Rapportage AI- & algoritmerisico's Nederland*. Den Haag: Autoriteit Persoonsgegevens.
- Autoriteit Persoonsgegevens (2023b). *Rapportage algoritmerisico's Nederland*. Den Haag: Autoriteit Persoonsgegevens.
- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052-1092.
- Bellamy, R.K., Dey, K., Hind, M., Hoffman, S.C., Houde, S., Kannan, K., & Nagar, S. (2018). Ai fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. arXiv preprint arXiv:1810.01943.
- Bogina, V., Hartman, A., Kuflik, T., & Shulner-Tal, A. (2022). Educating software and AI stakeholders about algorithmic fairness, accountability, transparency and ethics. *International Journal of Artificial Intelligence in Education*, 1-26.
- Bureau Common Ground., & it's public (2020). *Verkenning naar het Nederlandse onderwijslab Artificiële Intelligentie*. Amsterdam: Bureau Common Ground, it's public.
- Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2023). Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice*, 25(1), 51-75.
- Dam, B. van., Notermans, M., & Vlieg, M. (2022). *Stimuleren implementatie digitale leermiddelen in het primair en voortgezet onderwijs op basis van gedragsinzichten*. Nijmegen: D&B.
- Driessen, G. (red.). *De doelgroepen van het onderwijsachterstandenbeleid: ontwikkelingen in prestaties en het advies voortgezet onderwijs*. Nijmegen: ITS, Radboud Universiteit Nijmegen.
- Dondorp, L., & Pijpers, R. (2020). *Gelijke kansen en digitalisering? Goede toegang is niet genoeg*. Verkregen van: <https://www.kennisnet.nl/beleid-en-organisatie/gelijke-kansen-en-digitalisering-goede-toegang-is-niet-genoeg/>
- Eegdeman, I.M. (2023). *Enhancing Study Success in Dutch Vocational Education*. PhD-Thesis, Vrije Universiteit Amsterdam. <https://doi.org/10.5463/thesis.100>
- Eegdeman, I., Cornelisz, I., Meeter, M., & Van Klaveren, C. (2023). Identifying false positives when targeting students at risk of dropping out. *Education Economics*, 31(3), 313-325.
- Eegdeman, I., Cornelisz, I., Van Klaveren, C., & Meeter, M. (2022) Computer or teacher: Who predicts dropout best? *Frontiers in Education*, 7:976922.
- Eguchi, A., Okada, H., & Muto, Y. (2021). Contextualizing AI Education for K-12 Students to Enhance Their Learning of AI Literacy Through Culturally Responsive Approaches. *KI - Künstliche Intelligenz*, 35(2), 153–161.
- Fahd, K, Venkatraman, S, Miah, Md Shah Jahan, M, & Ahmed, Khandakar (2021). Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: a meta-analysis of literature. *Education and Information Technologies*, 27(3), 3743-3775.
- Hekkelman, B., Kattenberg, M., & Scheer, B. (2023). *Rechtvaardige algoritmes*. Den Haag: Centraal Planbureau.

- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S.B., Santos, O.C., Rodrigo, M.T., Cukurova, M., Bittencourt, I.I., & Koedinger, K.R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32, 504-526.
- Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development*, 68, 1961-1990.
- Inspectie van het Onderwijs (2024). *Peil.Digitale Geletterdheid Einde basisonderwijs 2021-2022*. Utrecht: Inspectie van het Onderwijs.
- Karssen, M., Krepel, A., Stronkhorst, E., Lourens, J.M.P., Bruck, S., Van Kessel, M. & Saab, N. (2023a). *Monitor digitalisering funderend onderwijs. Onderzoeksrapport Primair onderwijs MYRA 2023*. Amsterdam: Kohnstamm Instituut.
- Karssen, M., Krepel, A., Stronkhorst, E., Lourens, J.M.P., Bruck, S., Van Kessel, M. & Saab, N. (2023b). *Monitor digitalisering funderend onderwijs. Onderzoeksrapport Voortgezet (speciaal) onderwijs MYRA 2023*. Amsterdam: Kohnstamm Instituut.
- Kasman, M., & Valant, J. (2019). *The opportunities and risks of K-12 student placement algorithms*. Brookings.
- Kearns, M., & Roth, A. (2020). *Ethical Algorithm Design*. *ACM SIGecom Exchanges*, 18(1), 31–36.
- Liang, W., Yuksekogonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4. <https://doi.org/10.1016/j.patter.2023.100779>
- Mansoury, M., Abdollahpouri, H., Pechenizkiy, M., Mobasher, B., & Burke, R. (2020, October). Feedback loop and bias amplification in recommender systems. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 2145–2148).
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6), 1–35.
- Memarian, B., & Doleck, T. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI), and higher education: A systematic review. *Computers and Education: Artificial Intelligence*, 100152.
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (2021). *Impact Assessment Mensenrechten en Algoritmes*. Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2016). *Brief van de Minister en Staatssecretaris van Onderwijs, Cultuur en Wetenschap (31 oktober 2016), Tweede Kamer, vergaderjaar 2016-2017, 34 550 VIII, nr. 16*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2019). *Digitaliseringsagenda primair en voortgezet onderwijs*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- OECD (2018). *The resilience of students with an immigrant background: Factors that shape well-being*. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264292093-en>
- Onderwijsraad (2022). *Inzet van intelligente technologie. Een verkenning*. Den Haag: Onderwijsraad.
- Pangrazio, L., & Selwyn, N. (2021). Towards a school-based ‘critical data education.’ *Pedagogy, Culture & Society*, 29(3), 431–448.
- PBLQ (2023). *Naar een waardengedreven toepassing van AI in het mbo*. Den Haag: PBLQ.
- Pijpers, R., Bomas, E., Dondorp, L., & Ligthart, J. (2020). *Waarden wegen*. Zoetermeer: Kennisnet.
- Hekkelman, B., Kattenberg, M., & Scheer, B. (2023). *Rechtvaardige algoritmes*. Den Haag: Centraal Planbureau.

- Rathenau Instituut (2022a). *Algoritmes afwegen – Verkenning naar maatregelen ter bescherming van mensenrechten bij profilering in de uitvoering*. Den Haag: Rathenau Instituut.
- Rathenau Instituut (2022b). *Naar hoogwaardig digitaal onderwijs*. Den Haag: Rathenau Instituut.
- Streefkerk, J., & van Putten, R. (2020). *De verleiding van het algoritme. Maakbaarheidsdrang in de datasamenleving geïllustreerd, essay Staatsrechtconferentie 2020*. Utrecht: Universiteit Utrecht.
- SURF (2022). *Promises of AI in Education: Discussing the impact of AI systems in educational practices*. Verkregen van: <https://www.surf.nl/de-belofte-van-ai-in-het-onderwijs>
- Unesco (2021). *Recommendation on the Ethics of Artificial Intelligence*. Verkregen van: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Velander, J., Taiye, M. A., Otero, N., & Milrad, M. (2023). Artificial Intelligence in K-12 Education: eliciting and reflecting on Swedish teachers' understanding of AI and its implications for teaching & learning. *Education and Information Technologies*, 1-21.
- Versnellingsplan Onderwijsorganisatie met ICT (2021). *Referentiekader privacy en ethiek voor studiedata*. Verkregen van: <https://www.versnellingsplan.nl/Kennisbank/referentiekader-privacy-en-ethiek-voor-studiedata/>
- Vervliet, T. (2022). *Digitale werving & selectie. Algoritmegebruik in het werving- & selectieproces van werkgevers en het bewustzijn van risico's op uitsluiting en discriminatie. Eindrapport*. Amsterdam: SEO Economisch Onderzoek.
- Zhou, X., Van Brummelen, J., & Lin, P. (2020). Designing AI learning experiences for K-12: emerging works, future opportunities and a design framework. *arXiv preprint arXiv: 2009*.

